



1/78

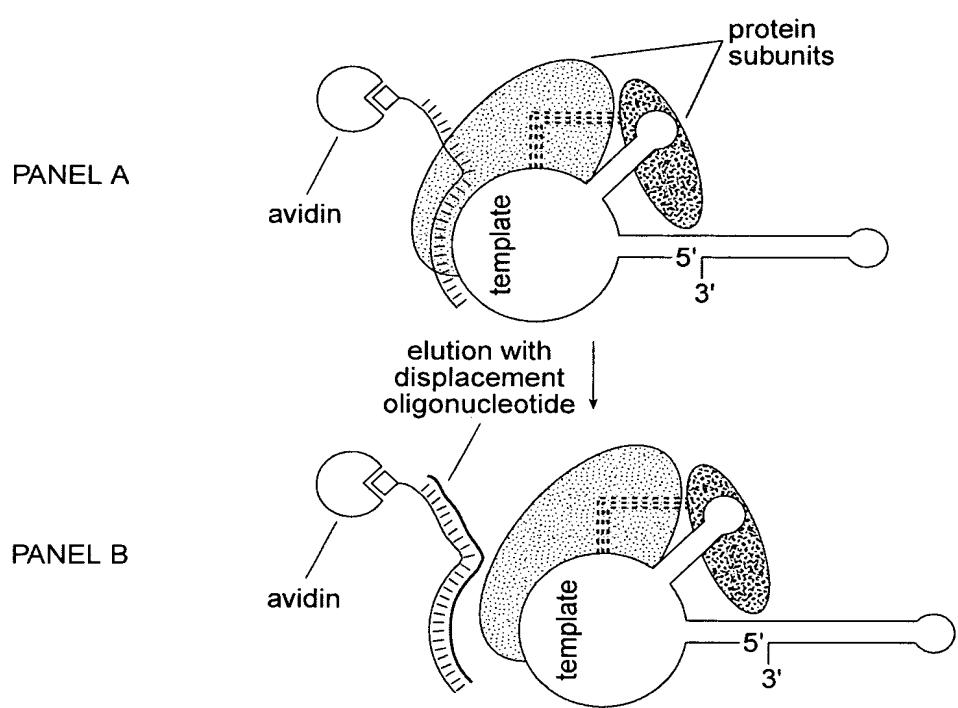


FIG. 1



2/78

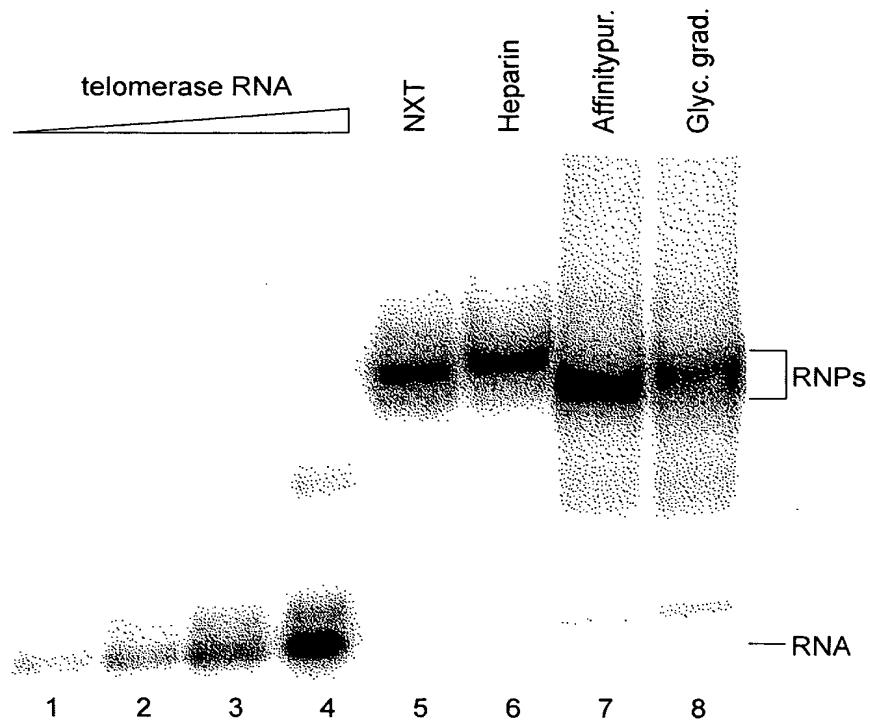


FIG. 2

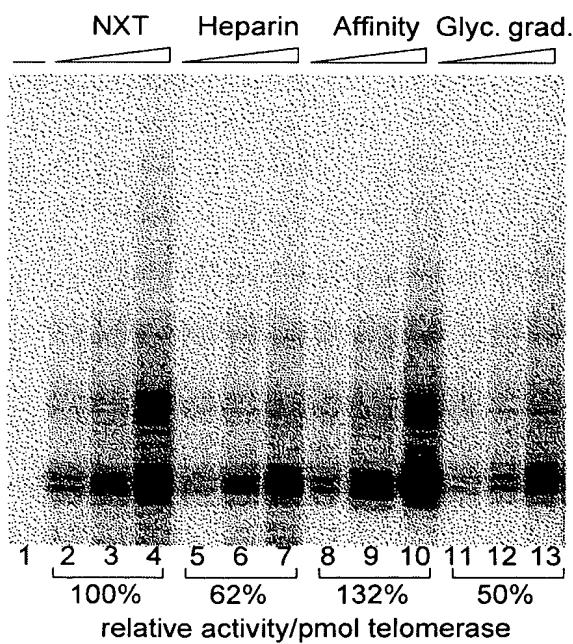


FIG. 3

+



3/78

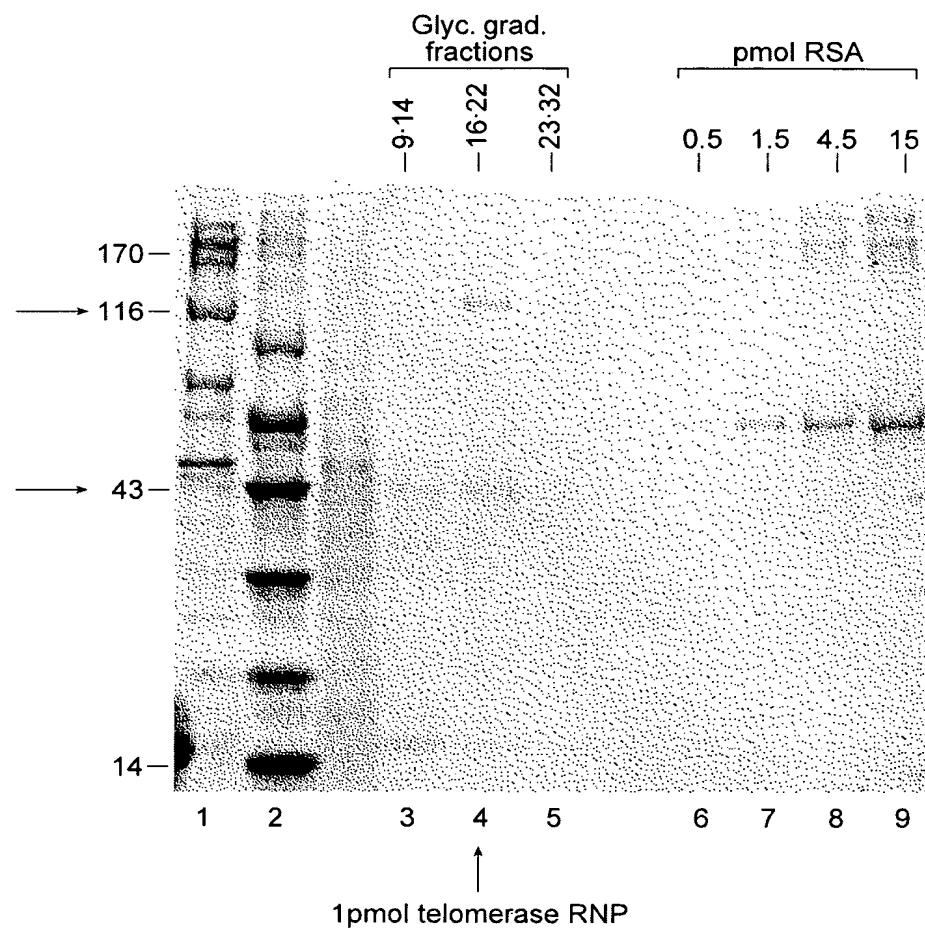


FIG. 4



4/78

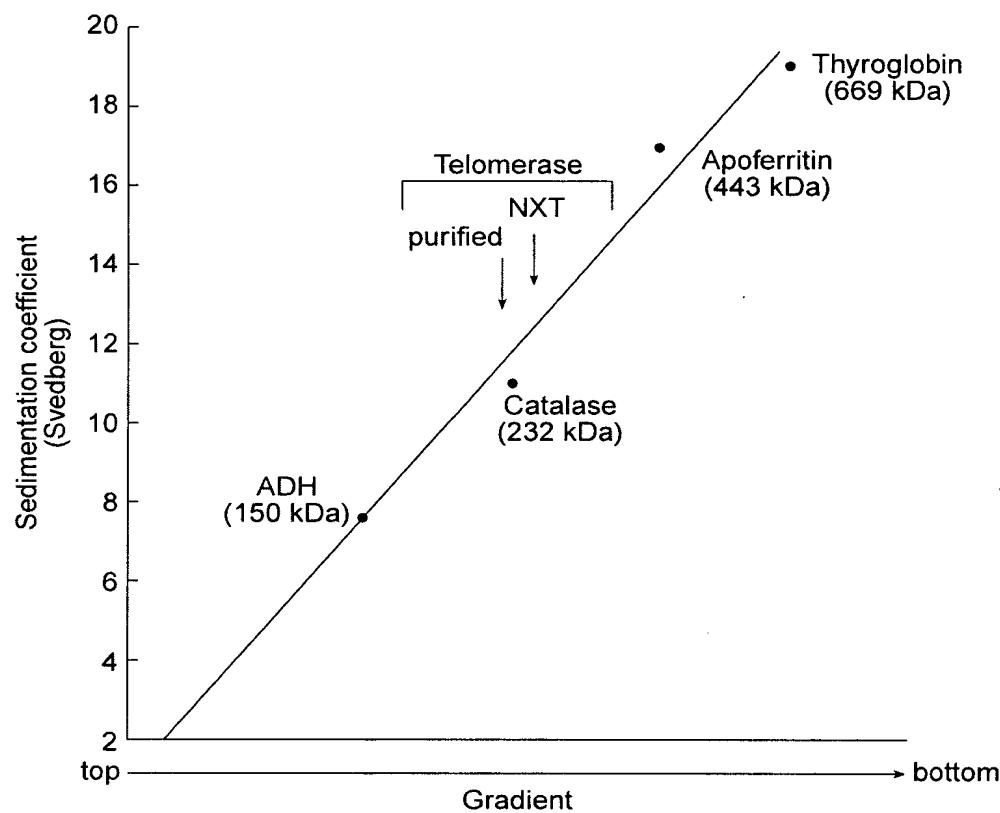


FIG. 5



5/78

Telomerase:

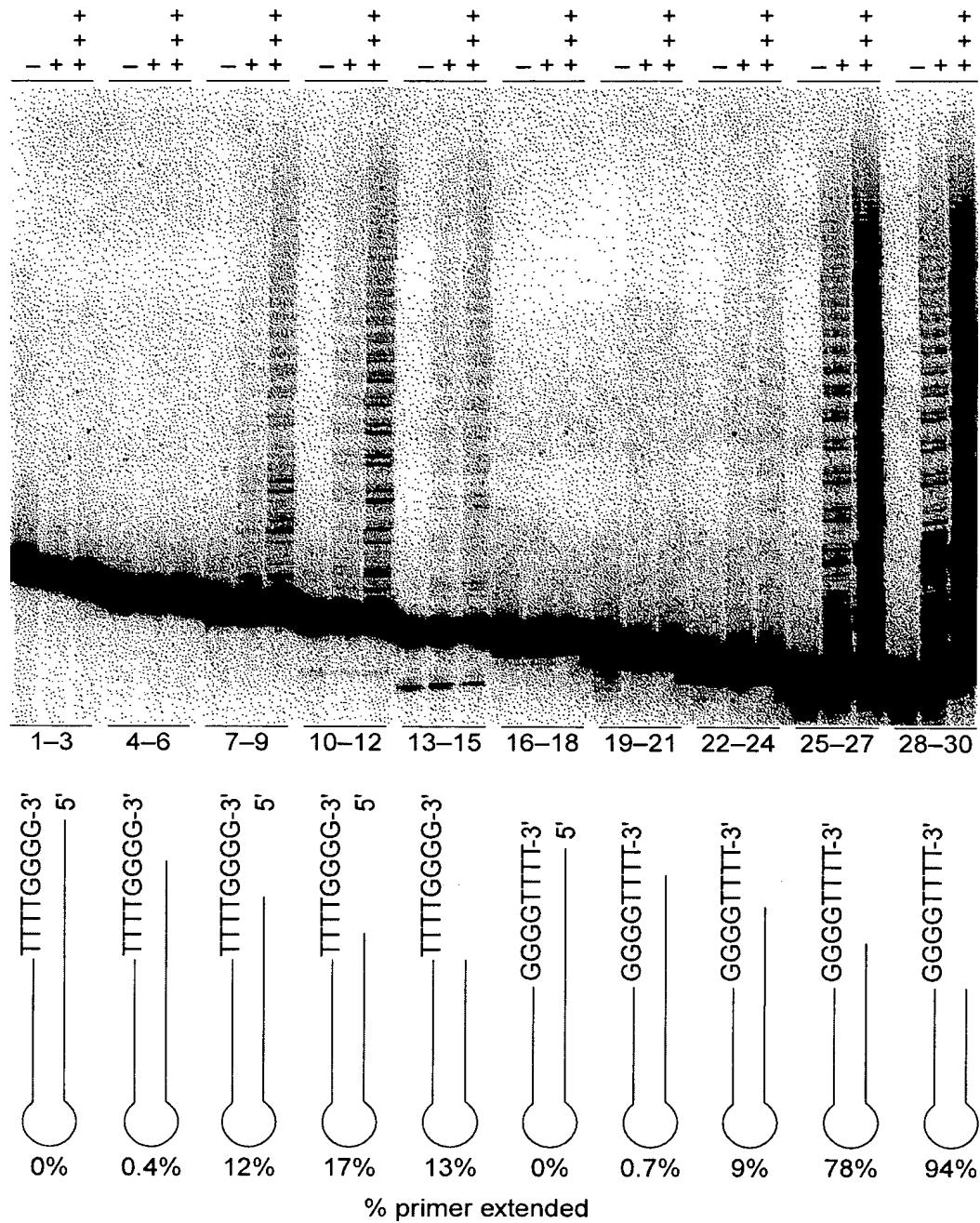


FIG. 6



6/78

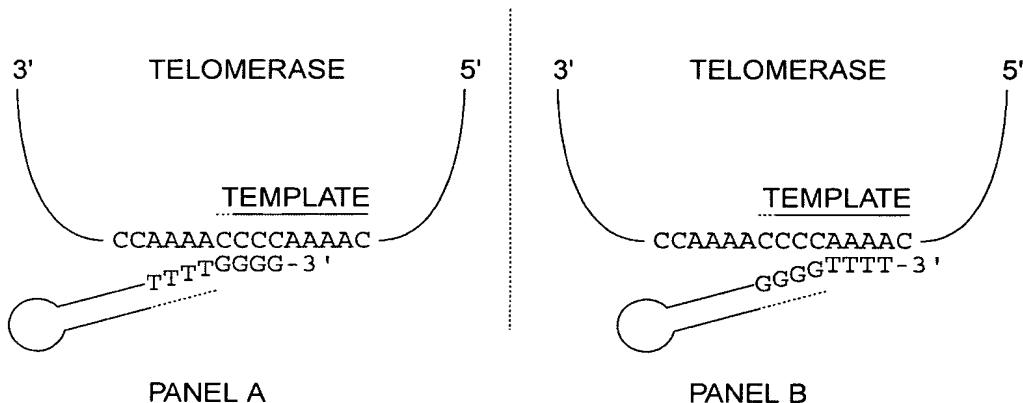


FIG. 7

1	CCCAAAACC	CCAAAACCCC	AAAACCCCTA	TAAAAAAAGA	AAAAAATTGAG
51	GTAGTTAGA	AATAAAATAT	TATTCCGCA	CAAATGGAGA	TGGATATTGA
101	TTGGATGAT	ATAGAAAATT	TACTTCCTAA	TACATTCAAC	AAGTATAGCA
151	GCTCTTGTAG	TGACAAGAAA	GGATGCAAAA	CATTGAAATC	TGGCTCGAAA
201	TCGGCTTCAT	TGACTATTCC	AAAGTTGCAA	AAACAATTAG	AGTTCTACTT
251	CTCGGATGCA	AATCTTATA	ACGATTCTT	CTTGAGAAAA	TTAGTTTAA
301	AAAGCGGAGA	GCAAAGAGTA	GAAATTGAAA	CATTACTAAT	GTAAATAAATAA
351	AATCAGGTAA	TGAGGATTAT	TCTATTTTT	AGATCACTTC	TTAAGGAGCA
401	TTATGGAGAA	AATTACTAA	TACTAAAAGG	TAAACAGTTT	GGATTATTTC
451	CCTAGCCAAC	AATGATGAGT	ATATCAAATT	CATATGAGAA	TGAGTCAAAG
501	GATCTCGATA	CATCAGACTT	ACCAAAGACA	AACTCGCTAT	AAAACGCAAG
551	AAAAAGTTTG	ATAATCGAAC	AGCAGAAGAA	CTTATTGCAT	TTACTATTG
601	TATGGGTTTT	ATTACAATTG	TTTTAGGTAT	CGACGGTGAA	CTCCCGAGTC
651	TTGAGACAAT	TGAAAAAGCT	GTTTACAACT	GAAGGAATCG	CAGTTCTGAA
701	AGTTCTGATG	TGTATGCCAT	TATTTTGTGA	ATTAATCTCA	AATATCTTAT
751	CTCAATTAA	TGGATAGCTA	TAGAAACAAA	CCAAATAAAC	CATGCAAGTT
801	TAATGGAATA	TACGTTAAT	CCTTTGGGAC	AAATGCACAC	TGAATTATAA
851	TTGGATTCTT	AAAGCATAGA	TACACAGAAT	GCTTTAGAGA	CTGATTTAGC
901	TTACAACAGA	TTACCTGTTT	TGATTACTCT	TGCTCATCTC	TTATATCTTT
951	AAAAGAAGCA	GGCGAAATGA	AAAGAAGACT	AAAGAAGAG	ATTTCAAAAT
1001	TTGTTGATTC	TTCTGTAACC	GGAAATTAAACA	ACAAGAATAT	TAGCAACGAA
1051	AAAGAAGAG	AGCTATCACA	ATCCTGATTC	TTAAAGATT	AAAAAATTCC
1101	AGGTAAAGAGA	GATACATTCA	TTAAAATTCA	TATATTATAG	TTTTTCATT
1151	CACAGCTGTT	ATTTTCTTTT	ATCTTAACAA	TATTTTTGA	TTAGCTGGAA
1201	GTAAAAAGTA	TCAAATAAGA	GAAGCGCTAG	ACTGAGGTAA	CTTAGCTTAT
1251	TCACATTCCAT	AGATCGACCT	TCATATATCC	AATACGATGA	TAAGGAAACAA
1301	GCAGTCATCC	GTTTTAAAAA	TAGTGCTATG	AGGACTAAAT	TTTTAGAGTC
1351	AAGAAATGGA	GCCGAAATCT	TAATCAAAAA	GAATTGCGTC	GATATTGCAA
1401	AAGAATCGAA	CTCTAAATCT	TTCGTTAATA	AGTATTACCA	ATCTTGATTG
1451	ATTGAAGAGA	TTGACGAGGC	AACTGCACAG	AAGATCATT	AAGAAATAAA
1501	GTAACATTAA	TTAATTAGAG	AATAAACTAA	ATTACTAATA	TAGAGATCAG
1551	CGATCTTCAA	TTGACGAAAT	AAAAGCTGAA	CTAAAGTTAG	ACAATAAAA
1601	ATACAAACCT	GGGTCAAAAT	ATTGAGGAAG	GAAGAGAAGA	CCAGTTAGCA
1651	AAAGAAAAAA	TAAGGCAATA	AATAAAATGA	GTACAGAACT	GAAGAAATAA
1701	AAGATTATT	TTTTCAATA	ATTATTGAA	AAGAGGGTT	TTGGGGTTTT
1751	GGGGTTTTGG	GG			

FIG. 11

+



7/78

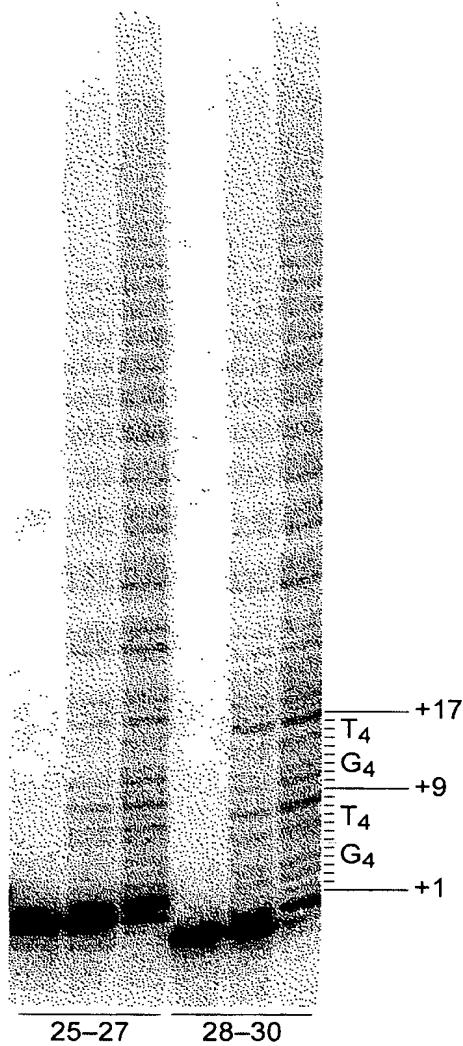


FIG. 8



1 AAAACCCCAA AACCCCAAAA CCCCTTTAG AGCCCTGCAG TTGGAAATAT  
 51 AACCTCAGTA TTAATAAGCT CAGATTTAA ATATTAATTA CAAAACCTAA  
 101 ATGGAGGTTG ATGTTGATAA TCAAGCTGAT AATCATGGCA TTCACTCAGC  
 151 TCTTAAGACT TGTGAAGAAA TTAAAGAAGC TAAAACGTTG TACTCTTGG  
 201 TCCAGAAAGT TATTAGATGA AGAAATCAAT CTCAAAGTCA TTATAAAGAT  
 251 TTAGAAGATA TTAAAATATT TGCGCAGACAA AATATTGTTG CTACTCCACG  
 301 AGACTATAAT GAAGAAGATT TTAAAGTTAT TGCAAGAAAA GAAGTATTTT  
 351 CAACTGGACT AATGATCGAA CTTATTGACA AATGCTTAGT TGAACCTTCTT  
 401 TCATCAAGCG ATGTTTCAGA TAGACAAAAA CTTCAATGAT TTGGATTTCA  
 451 ACTTAAGGGA AATCAATTAG CAAAGACCCA TTTATTAACA GCTCTTCAA  
 501 CTCAAAAGCA GTATTCTTT CAAGACGAAT GGAACCAAGT TAGAGCAATG  
 551 ATTGGAAATG AGCTCTTCCG ACATCTCTAC ACTAAATATT TAATATTCCA  
 601 GCGAACTTCT GAAGGAACTC TTGTTCAATT TTGCGGGAAAT AACGTTTTG  
 651 ATCATTGAA AGTCAACGAT AAGTTTGACA AAAAGCAAAA AGGTGGAGCA  
 701 GCAGACATGA ATGAACCTCG ATGTTGATCA ACCTGCAAAT ACAATGTC  
 751 GAATGAGAAA GATCACTTT TCAACAAACAT CAACGTGCCG AATTGGAATA  
 801 ATATGAAATC AAGAACCCAGA ATATTTTATT GCACTCATT TAATAGAAAT  
 851 AACCAATTCT TCAAAAAGCA TGAGTTTGTG AGTAACAAAAA ACAATATTTC  
 901 AGCGATGGAC AGAGCTCAGA CGATATTAC GAATATATTAG ATGTTAATA  
 951 GAATTAGAAA GAAGCTAAAAA GATAAGGTTA TCGAAAAAAAT TGCCTACATG  
 1001 CTTGAGAAAG TCAAAGATT TAACTTCAAC TACTATTTAA CAAAATCTT  
 1051 TCCTCTTCCA GAAAATTGGC GGGAACGGAA ACAAAAAATC GAAAACCTTGA  
 1101 TAAATAAAAC TAGAGAAGAA AAGTCGAAGT ACTATGAAGA GCTGTTAGC  
 1151 TACACAACTG ATAATAAAATG CGTCACACAA TTTATTAATG AATTTTCTA  
 1201 CAATATACTC CCCAAAGACT TTTTGACTGG AAGAAACCGT AAGAATTTC  
 1251 AAAAGAAAGT TAAGAAATAT GTGGAACCTAA ACAAGCATGA ACTCATTAC  
 1301 AAAAACTTAT TGCTTGAGAA GATCAATACA AGAGAAATAT CATGGATGCA  
 1351 GGTTGAGACC TCTGCAAAGC ATTTTTATTA TTTTGATCAC GAAAACATCT  
 1401 ACGTCTTATG GAAATTGCTC CGATGGATAT TCGAGGATCT CGTCGCTCG  
 1451 CTGATTAGAT GATTTTCTA TGTCACCGAG CAACAGAAAA GTTACTCCAA  
 1501 AACCTATTAC TACAGAAAGA ATATTTGGGCG CGTCATTATG AAAATGTC  
 1551 TCGCAGACTT AAAGAAGGAA ACGCTTGCTG AGGTCCAAGA AAAAGAGGTT  
 1601 GAAGAATGGA AAAAGTCGCT TGGATTTGCA CCTGGAAAAC TCAGACTAAT  
 1651 ACCGAAGAAA ACTACTTTCC GTCCAATTAT GACTTTCAAT AAGAAGATTG  
 1701 TAAATCTCAGA CCGGAAGACT ACAAAATTAA CTACAAATAC GAAGTTATTG  
 1751 AACTCTCACT TAATGCTTAA GACATTGAG AATAGAATGT TTAAAGATCC  
 1801 TTTGGATTC GCTGTTTTA ACTATGATGA TGTAAATGAAA AAGTATGAGG  
 1851 AGTTTGTGTTG CAAATGGAAG CAAGTTGGAC AACCCAAACT CTTCTTGCA  
 1901 ACTATGGATA TCGAAAAGTG ATATGATAGT GTAAACAGAG AAAAATATC  
 1951 AACATTCTA AAAACTACTA ATTACTTTC TTCAAGATTTG TGGATTATGA  
 2001 CTGCACAAAT TCTAAAGAGA AAGAATAACA TAGTTATCGA TTGAAAAAC  
 2051 TTTAGAAAGA AAGAAATGAA AGATTATTT AGACAGAAAT TCCAGAAGAT  
 2101 TGCACCTGAA GGAGGACAAT ATCCAACCTT ATTCAAGTGT CTTGAAAATG  
 2151 AACAAAATGA CTTAAATGCA AAGAAAACAT TAATTGTTGA AGCAAAGCAA  
 2201 AGAAATTATT TTAAGAAAGA TAACTTACTT CAACCAGTCA TTAATATTG  
 2251 CCAATATAAT TACATTAAC TTAATGGAA GTTTTATAAA CAAACAAAAG  
 2301 GAATTCCCTCA AGGTCTTGA GTTTCATCAA TTTTGTATC ATTTTATTAT  
 2351 GCAACATTAG AGGAAAGCTC CTTAGGATTC CTTAGAGATG AATCAATGAA

FIG. 9A



9/78

2401 CCCTGAAAAT CCAAATGTTA ATCTTCTAAT GAGACTTACA GATGACTATC  
 2451 TTTTGATTAC AACTCAAGAG ATAATGCAG TATTGTTTAT TGAGAAACTT  
 2501 ATAAACGTA GTCGTGAAAA TGGATTTAAA TTCAATATGA AGAAACTACA  
 2551 GACTAGTTT CCATTAAGTC CAAGCAAATT TGCAAAATAC GGAATGGATA  
 2601 GTGTTGAGGA GCAAAATATT GTTCAAGATT ACTGCGATTG GATTGGCATC  
 2651 TCAATTGATA TGAAAATCT TGCTTTAATG CCAAATATTA ACTTGAGAAT  
 2701 AGAAGGAATT CTGTGTACAC TCAATCTAAA CATGCAAACA AAGAAAGCAT  
 2751 CAATGTGGCT CAAGAAGAAA CTAAAGTCGT TTTAATGAA TAACATTACC  
 2801 CATTATTTA GAAAGACGAT TACAACCGAA GACTTTGCGA ATAAAACCTCT  
 2851 CAACAAGTTA TTTATATCAG GCGGTTACAA ATACATGCAA TGAGCCAAAG  
 2901 AATACAAGGA CCACTTAAAG AAGAACTTAG CTATGAGCAG TATGATCGAC  
 2951 TTAGAGGTAT CTAAAATTAT ATACTCTGTA ACCAGAGCAT TCTTTAAATA  
 3001 CCTTGTGTGC AATATTAAGG ATACAATTTC TGGAGAGGAG CATTATCCAG  
 3051 ACTTTTCCCT TAGCACACTG AAGCACTTTA TTGAAATATT CAGCACAAAA  
 3101 AAGTACATT TCAACAGAGT TTGCATGATC CTCAGGCAA AAGAAGCAAA  
 3151 GCTAAAAAGT GACCAATGTC AATCTCTAAT TCAATATGAT GCATAGTCGA  
 3201 CTATTCTAAC TTATTTGGA AAGTTAATT TCAATTTTG TCTTATATAC  
 3251 TGGGGTTTTG GGGTTTGGG GTTTTGGGG

FIG. 9B

1 MEVDVDNQAD NHGIHSALKT CEEIKEAKTL YSWIQKVIRC RNQSQSHYKD  
 51 LEDIKIFQAQT NIVATPRDYN EEDFKVIARK EVFSTGLMIE LIDKCLVELL  
 101 SSSDVSDRQK LQCFGFQLKG NQLAKTHLLT ALSTQKQYFF QDEWNQVRAM  
 151 IGNELFRHLY TKYLIFQRTS EGTLVQFCGN NVFDHLKVND KFDKKQKGGA  
 201 ADMNEPRCCS TCKYNVKNEK DHFLNNINVP NWNNMKSRTTR IFYCTHFNRN  
 251 NQFFKKHEFV SNKNNISAMD RAQTIFTNIF RFNRIRKKLK DKVIEKIAM  
 301 LEKVKDFNFN YYLTKSCPLP ENWRERKQKI ENLINKTREE KSKYYEELFS  
 351 YTTDNKCVTQ FINEFFYNIL PKDFLTGRNR KNFQKKVKKY VELNKHели  
 401 KNLLEKINT REISWMQVET SAKHFYYFDH ENIYVLWKL RWIFEDLVVS  
 451 LIRCFFYVTE QQKSYSKTY YRKNIWDVIM KMSIADLKKE TLAEVQEKEV  
 501 EEWKKSLGFA PGKLRLIPKK TTFRPIMTFN KKIVNSDRKT TKLTTNTKLL  
 551 NSHMLMLKTLK NRMFKDPFGF AVFNYDDVMK KYEEFVCKWK QVGQPKLFFA  
 601 TMDIEKCYDS VNREKLSTFL KTTKLLSSDF WIMTAQILKR KNNIVIDSKN  
 651 FRKKEMKDYF RQKFQKIALE GGQYPTLFSV LENEQNDLNA KKTLLIVEAKQ  
 701 RNYFKKDNLQ QPVINICQYN YINFNGKFYK QTKGIPQGLC VSSILSSFYY  
 751 ATLEESSLGF LRDESMNPEN PNVNLLNMRLT DDYLLITTQE NNAVLFIEKL  
 801 INVSRENGFK FNMKKLQTSF PLSPSKFAKY GMDSVEEQNI VQDYCDWIGI  
 851 SIDMKTALM PNINLRIEGI LCTLNLMQT KKASMWLKKK LKSFLMNNIT  
 901 HYFRKTITTE DFANKTLNKL FISGGYKYMQ CAKEYKDHFK KNLAMSSMID  
 951 LEVSKIIYSV TRAFFKYLVC NIKDTIFGEE HYPDFFLSTL KHFIEIFSTK  
 1001 KYIFNRVCMI LKAKEAKLKS DQCQSLIQYD A

FIG. 10



10/78

1 CCCCCAAAACCCAAAACCCAAAACCCCTATAAAAAAGAAAAATTGAGGTAGTTAGA 60  
 GGGGTTTGGGGTTTGGGGTTTGGGGATATTTTTCTTTTAACCCATCAAATCT  
  
 a P Q N P K T P K P L \* K K K K L R \* F R -  
 b P K T P K P Q N P Y K K R K N C G S L E -  
 c P K P Q N P K T P I K K E K I E V V \* K -  
  
 AATAAAAATTATTATTCCCGCACAAATGGAGATGGATATTGATTTGGATGATATAGAAAATT 120  
 61 TTATTTATAATAAGGGCGTGTACCTACCTATAACTAACCTACTATATCTTTAA  
  
 a N K I L F P H K W R W I L I W M I \* K I -  
 b I K Y Y S R T N G D G Y C F G C Y R K F -  
 c \* N I I P A Q M E M D I D L D D I E N L -  
  
 TACTTCCTAATACATTCAACAAGTATAGCAGCTTGTAGTGACAAGAAAGGATGCAAAA 180  
 121 ATGAAGGATTATGTAAGTTGTCATATCGTCGAGAACATCACTGTTCTTCCTACGTTT  
  
 a Y F L I H S T S I A A L V V T R K D A K -  
 b T S \* Y I Q Q V \* Q L L \* \* Q E R M Q N -  
 c L P N T F N K Y S S S C S D K K G C K T -  
  
 CATTGAAATCTGGCTCGAAATCGCCTTCATTGACTATTCAAAGTTGCAAAAACAATTAG 240  
 181 GTAACCTTAGACCGAGCTTAGCGGAAGTAACTGATAAGGTTCAACGTTTGTAAATC  
  
 a H C N L A R N R L H C L F Q S C K N N \* -  
 b I E I W L E I A F I D Y S K V A K T I R -  
 c L K S G S K S P S L T I P K L Q K Q L E -  
  
 AGTTCTACTTCTCGGATGCAAATCTTATAACGATTCTTCTTGAGAAAATTAGTTTAA 300  
 241 TCAAGATGAAGAGCCTACGTTAGAAATATTGCTAAGAAAGAACTCTTTAATCAAAATT  
  
 a S S T S R M Q I F I T I L S C E N \* F \* -  
 b V L L L G C K S L \* R F F L E K I S F K -  
 c F Y F S D A N L Y N D S F L R K L V L K -  
  
 AAAGCGGAGAGCAAAGAGTAGAAACATTACTAATGTTAAATAAAATCAGGTAA 360  
 301 TTTCGCCTCTGTTCTCATCTTAACCTTGTAAATGATTACAAATTATTTAGTCCATT  
  
 a K A E S K E \* K L K H Y \* C L N K I R \* -  
 b K R R A K S R N C N I T N V \* I K S G N -  
 c S G E Q R V E I E T L L M F K \* N Q V M -  
  
 TGAGGATTATTCTATTTTAGATCACTCTTAAGGAGCATTATGGAGAAAATTACTTAA 420  
 361 ACTCCTAATAAGATAAAAATCTAGTGAAGAATTCCCTCGTAATACCTCTTTAATGAATT  
  
 a C G L F Y F L D H F L R S I M E K I T \* -  
 b E D Y S I F \* I T S \* G A L W R K L L N -  
 c R I I L F F R S L L K E H Y G E N Y L I -

FIG. 12A



+

11/78

TACTAAAAGGTAAACAGTTGGATTATTCCTAGCCAACAATGATGAGTATTTAAATT  
 421 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 480  
 ATGATTTCCATTTGTCAAACCTAATAAAGGGATCGGTTACTACTCATATAATTAA  
  
 a Y \* K V N S L D Y F P S Q Q C C C V Y \* I -  
 b T K R \* T V W I I S L A N N D E Y I K F -  
 c L K G K Q F G L F P \* P T M M S I L N S -  
  
 CATATGAGAATGAGTCAAAGGATCTGATACATCAGACTTACCAAGACAAACTCGCTAT  
 481 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 540  
 GTATACTCTTACTCAGTTCTAGAGCTATGTAGTCTGAATGGTTCTGAGCGATA  
  
 a H M R M S Q R I S I H Q T Y Q R Q T R Y -  
 b I C E C V K G S R Y I R L T K D K L A I -  
 c Y E N E S K D L D T S D L P K T N S L \* -  
  
 AAAACGCAAGAAAAAGTTGATAATCGAACAGCAGAAGAACTTATTGCATTTACTATTG  
 541 -----+-----+-----+-----+-----+-----+-----+-----+ 600  
 TTTTGCGTTCTTTCAAACATTAGCTGTGCTCTTGAATAACGTAATGATAAGC  
  
 a K T Q E K V C \* S N S R R T Y C I Y Y S -  
 b K R K K K F D N R T A E E L I A F T I R -  
 c N A R K S L I I E Q Q K N L L H L L F V -  
  
 TATGGGTTTATTACAATTGTTAGGTATCGACGGTGAACCTCCGAGTCTTGAGACAAT  
 601 -----+-----+-----+-----+-----+-----+-----+ 660  
 ATACCCAAAATAATGTTAACAAATCCATAGCTGCCACTTGAGGGCTCAGAACTCTGTTA  
  
 a Y G F Y Y N C F R Y R R C T P E S C D N -  
 b M G F I T I V L G I D G E L P S L E T I -  
 c W V L L Q L F \* V S T V N S R V L R Q L -  
  
 TGAAAAGCTGTTACAACCTGAGGAATCGCAGTTCTGAAAGTTCTGATGTGTATGCCAT  
 661 -----+-----+-----+-----+-----+-----+-----+ 720  
 ACTTTTCGACAAATGTTGACTTCTTAGCGTCAAGACTTCAAGACTACACATACGGTA  
  
 a C K S C L Q L K E S Q F C K F \* C V C H -  
 b E K A V Y N \* R N R S S E S S D V Y A I -  
 c K K L F T T E G I A V L K V L M C M P L -  
  
 TATTTGTGAATTAAATCTCAAATATCTTATCTCAATTAAATGGATAGCTATAGAAACAAA  
 721 -----+-----+-----+-----+-----+-----+-----+ 780  
 ATAAAACACTTAATTAGAGTTATAGAAATAGAGTTAACCTATCGATATCTTGT  
  
 a Y F V N \* S Q I S Y L N L M D S Y R N K -  
 b I L \* I N L K Y L I S I \* W I A I E T N -  
 c F C E L I S N I L S Q F N G \* L \* K Q T -  
  
 CCAAATAAACCATGCAAGTTAATGGAATATACGTTAACCTTGGGACAAATGCACAC  
 781 -----+-----+-----+-----+-----+-----+-----+ 840  
 GGTTTATTGGTACGTTCAAATTACCTTATATGCAATTAGAAACCTGTTACGTGT  
  
 a P N K P C K F N G I Y V K S F G T N A H -  
 b Q I N H A S L M E Y T L N P L G Q M H T -  
 c K \* T M Q V \* W N I R \* I L W D K C T L -  
  
 TGAATTATGGATTCTAAAGCATAGATACACAGAATGCTTAGAGACTGATTAGC  
 841 -----+-----+-----+-----+-----+-----+-----+ 900  
 ACTTAAATATAACCTAACGAAATTGCTATATGCTTACGAAATCTGACTAAATCG  
  
 a C I Y I G F L K H R Y T E C F R D C F S -  
 b E F I L D S \* S I D T Q N A L E T D L A -  
 c N L Y W I L K A \* I H R M L \* R L I \* L -

FIG. 12B

+



12/78

TTACAAACAGATTACCTGTTGATTACTCTGCTCATCTCTTATATCTTAAAGAAGCA  
 901 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 960  
 AATGTTGTCTAATGGACAAACTAATGAGAACGAGTAGAGAAATAGAAATTTCTCGT

a L Q Q I T C F D Y S C S S L I S L K E A -  
 b Y N R L P V L I T L A H L L Y L \* K K Q -  
 c T T D Y L F \* L L L L I S Y I F K R S R -

GGCGAAATGAAAAGAAGACTAAAGAAAGAGATTCAAAATTGTTGATTCTCTGTAACC  
 961 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 1020  
 CCGCTTACTTTCTGATTCTCTAAAGTTAAACAACTAAGAACATTGG

a G E M K R R L K K E I S K F V D S S V T -  
 b A K \* K E D \* R K R F Q N L L I L L \* P -  
 c R N E K K T K E R D F K I C \* F F C N R -

GGAATTAAACAACAAGAATATTAGCAACGAAAAAGAAGAACAGACTATCACAAATCCTGATT  
 1021 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 1080  
 CCTTAATTGTTGTTCTTATAATCGTTGCTTTCTCTCGATAGTGTAGGACTAAG

a G I N N K N I S N E K E E E L S Q S \* F -  
 b E L T T R I L A T K K K K S Y H N P D S -  
 c N \* Q Q E Y \* Q R K R R R A I T I L I L -

TTAAAGATTCAAAATTCCAGGTAAGAGAGATACTTCAATTAAATTCAATTATATTAG  
 1081 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 1140  
 AATTCTAAAGTTTAAGGTCCATTCTCTATGTAAGTAATTAAAGTATATAATTC

a L K I S K I P G K R D T F I K I H I L \* -  
 b \* R F Q K F Q V R E I H S L K F I Y Y S -  
 c K D F K N S R \* E R Y I H \* N S Y I I V -

TTTTTCATTCACAGCTGTTATTTCTTATCTTAAACAATTGTTGATTAGCTGGAA  
 1141 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 1200  
 AAAAGTAAAGTGTGACAAATAAGAAAATAGAATTGTTATAAAAACATAATCGACCTT

a F F I S Q L L F S F I L T I F F D \* L E -  
 b F S F H S C Y F L L S \* Q Y F L I S W K -  
 c F H F T A V I F F Y L N N I F \* L A G S -

GTAAAAGTATCAAATAAGAGAACAGCGCTAGACTGAGGTAACCTAGCTTATTCACATT  
 1201 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 1260  
 CATTTCATAGTTATTCTCTCGCGACTGACTCCATTGAATCGAATAAGTGTAGTA

a V K S I K \* E K R \* T E V T \* L I H I H -  
 b \* K V S N K R S A R L R \* L S L F T F I -  
 c K K Y Q I R E A L D \* G N L A Y S H S \* -

AGATCGACCTTCATATATCCAATACGATGATAAGGAAACAGCAGTCATCCGTTTAAAAA  
 1261 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 1320  
 TCTAGCTGGAAGTATAGGTTATGCTACTATTCTTGTGTCAGTAGGCAAATT  

a R S T F I Y P I R \* \* G N S S H P F \* K -  
 b D R P S Y I Q Y D D K E T A V I R F K N -  
 c I D L H I S N T M I R K Q Q S S V L K I -

TAGTGCTATGAGGACTAAATTAGAGTCAGGAAATGGAGCCGAAATCTTAATCAAAA  
 1321 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 1380  
 ATCACGATACTCCTGATTAAAGTCTAGTTACCTCGGCTTGAATTAGTTT

a \* C Y E D \* I F R V K K W S R N L N Q K -  
 b S A M R T K F L E S R N G A E I L I K K -  
 c V L \* G L N F \* S Q E M E P K S \* S K R -

FIG. 12C



13/78

GAATTGCGTCGATATTGCAAAAGAACATCGAACTCTAAATCTTCGTTAATAAGTATTACCA  
 1381 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 1440  
 CTTAACGCGACTATAACGTTCTTAGCTGAGATTAGAAAGCAATTATTCATAATGGT

a E L R R Y C K R I E L \* I F R \* \* V L P -  
 b N C V D I A K E S N S K S F V N K Y Y Q -  
 c I A S I L Q K N R T L N L S L I S I T N -

ATCTTGATTGATTGAAGAGATTGACGAGGCAACTGCACAGAAGATCATTAAAGAAATAAA  
 1441 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 1500  
 TAGAACTAACTAACTCTACTGCTCCGTTGACGTGCTTAGTAATTCTTATTT

a I L I D C R D \* R G N C T E D H \* R N K -  
 b S C L I E E I D E A T A Q K I I K E I K -  
 c L D C L K R L T R Q L H R R S L K K \* S -

GTAACTTTATTAAATTAGAGAATAAAACTAAATTACTAATATAGAGATCAGCGATCTCAA  
 1501 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 1560  
 CATTGAAAATAATTAATCTCTTATTGATTAAATGATTATCTCTAGTCGCTAGAAGTT

a V T F I N \* R I N \* I T N I E I S D L Q -  
 b \* L L L I R E \* T K L L I \* R S A I F N -  
 c N F Y \* L E N K L N Y \* Y R D Q R S S I -

TTGACGAAATAAAAGCTGAACAAAGTTAGACAATAAAACCTGGTCAAAAT  
 1561 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 1620  
 AACTGCTTATTTCGACTTGATTCAATCTGTTATTGTTATGTTGGAACAGTTTA

a L T K \* K L N \* S \* T I K N T N L G Q N -  
 b C R N K S C T K V R Q \* K I Q T L V K I -  
 c D E I K A E L K L D N K K Y K P W S K Y -

ATTGAGGAAGGAAAAGAACAGCAGTTAGCAAAAGAAAAATAAGGCAATAAAATAAAATGA  
 1621 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 1680  
 TAACTCCTCCTTTCTGGTCAATCGTTTTTATTCCGTTATTACT

a I E E G K E D Q L A K E K I R Q \* I K C -  
 b L R K E K K T S \* Q K K K \* G N K \* N E -  
 c C G R K R R P V S K R K N K A I N K M S -

GTACAGAAAGTGAAGAAATAAAAGATTATTTCGTTCAATAATTATTGAAAAGAGGGGTT  
 1681 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 1740  
 CATGTCTTCACTCTTATTCTAAATAAAAGTTATTAAATAACTTTCTCCCCAA

a V Q K C R N K R F I F F N N L L K R G V -  
 b Y R S E E I K D L F F S I I Y C K E G F -  
 c T E V K K \* K I Y F F Q \* F I E K R G F -

TTGGGGTTTGGGGTTTGGGG  
 1741 -----+-----+--- 1762  
 AACCCCCAAAACCCCAAAACCCC

a L G F W G F G -  
 b W G F G V L G -  
 c G V L G F W -

FIG. 12D



14/78

2 EVDVQNQADNHGIHSALKTCEEIKEAKTLYSWIQKVIRCRNQSQSHYKDL 51  
 19 ELELEMQENQNDIQVRVK....IDDPKQY..LVNVTAACLLQEGSYYQDK 62  
 52 EDIKIFAQTNIVATPRDYNEEDFKVIARKEVF.STGLMIELIDKCLVELL 100  
 63 DERRYIITKALL....EVAESDPFECQLAVYIRNELYIRTTNYIVAF. 107  
 101 SSSDVSDRQKLQCFGQLKGNQLAKTHLLTALSTQKQYFFQDEWNQVRAM 150  
 108 .....CVVHKNTQPFIEKYFNKAVLLPNDLLEVCEFAQVLYI 144  
 151 IGNELFRHLYTKYLIFQRTSEGTLVQFCGNNVFDHLKVNDKFDKKQKGGA 200  
 145 FDATEFKNLY.....LDRILSQDIRKELTRKCLQRCVRSKF 181  
 201 ADMNE...PRCCSTCKYNVKNEKDHFLLNNINVPNWNMKSRTIFYCTHF 247  
 182 SEFNEYQLGKYCTES..QRKKTMFRYLSVTNKQKWDQTKKK..... 220  
 248 NRNNQFFKKHEFVSNKNNISAMDRAQTIFTNIFRFNRIRKKLKDVKIEKI 297  
 221 .RKENNLLTKLQAIKESEDKSKEKG....DIMNVEDAIKALKPAVMKKI 264  
 298 AYMLEKVKDFNFNYYLTKSCPLPENWRERKQKIEALKTREESKYYEE 347  
 265 AKRQNAMEK.....KHMKAPKIPNSTLESKYLTFKD 294  
 348 LFSYTTDNKCVTQFINEFFYNILPKDFLTGRNRKNFQKKVKKYVELNKH 397  
 295 LIKFCHISEP.....KERVYKILGKKYPKTEEEYKAAFGDSASAPFN.PE 338  
 398 LIHKNLLLEKINTREISWMQVETSAKHFYFYDHENIYVWLWKLRLWIFEDL 447  
 339 LAGKRMKIEISKTWENELSAKGNTAEVWDNLISSNQLPYMAMLRNLSN.. 386  
 448 VVSLIRCFYVTEQQKSYSKTYYRKNIWDVIMKMSIADLKETLAEVQE 497  
 387 .....ILKAGVSD..... 394  
 498 KEVEEWKKSLGFAPGKLRЛИPKTTFRPIMTFNKKIVNSDRKTTKLTNT 547  
 395 .....TTHS 398  
 548 KLLNSHMLKTLKNRMFKDPFGFAVFNYDDVMKKYEEFVCKWKQVGQPKL 597  
 399 IVINK.....ICEPKAVENSKM 415  
 598 FFATMDIEKCYDSVNREKLSTFLKTTKLLSSDFWIMTAQILKRKNNIVID 647  
 416 F..PLQFFSAIEAVN.EAVTKGFKAKK...RENMLKGQIEAVKE..VVE 457  
 648 SKNFRKKEMKDYFRQKFQKIALEGGQYPTLFSVLENEQNDLNAKKTIVE 697  
 458 KTDEEKKDM.....ELEQTEEGEFVKVNEGIGKQYINSIELAIK 496  
 698 AKQRNYFKKDNLQPVINICQYNYINFNGKFYKQTKGIPQGLCVSSILSS 747  
 497 IAVNKNLDEIKGHTAIFSDVSGSMSTSMSGGAKKYGSVRTCLECALVLGL 546  
 748 FYYATLEESSLGFLRDESMNPENPNVNLLMRLTDDYLLITTQENNAVLFI 797  
 547 MVKQRCEKSSFYIFSSPSSQCNKCYLEVNL..... 576

FIG. 13A

+



15/78

798 EKLINVSRENGFKFNMKK. LQTSFPLSPSKFAKYGMDSVEEQNIVQDYCD 846  
 577 .....PGDELRPSMQKLLQEKGKLGGS. TDFPYECIDEWTKNKTHVD 617  
 847 WIGISIDMKTALALMPNINLRIEGILCTLNLNMQTKKASMWLKKKLKSFLM 896  
 618 NIVILSDMMIAEGYSDINVRGSSIVNSI.....KKYKDEVN 653  
 897 NNITHYFRKTITTEDFANKTLNKLFISSGGYKYMQCAKEYKD. HFKNLAM 945  
 654 PNIKIF...AVDLEGYG.....KCLNLGDEFNENNYIKIFGM 687  
 946 SSMIDLEVSKIIYSVTRAFFKYLCVNIKDTIFGEEHYPDFFLSTLKHFIE 995  
 688 SDSI.....LKFISAKQGGA.....NMVE 706  
 996 IFSTKKYIFNRVC 1008  
 707 VI..KNFALQKIG 717

FIG. 13B

132 LSTQKQYFFQDEWNQVRAMIGNEL. FRHLYTKYLIFQRTSE..GTLVQFC 178  
 1 MSRRNQ.....KKPQAPIGNETNLDFVLQNLEVYKSQIEHYKTQQQI 43  
 179 GNNVFDHLKVNDKFDKKQKGGAADMNEPRCCSTCKYNVNEKDHFLNNIN 228  
 44 KEEDLKLLKFKNQDQDGNSGNDDDEE.....NNSNKQQELLRRVN 84  
 229 VPKWNNMKSRTTRIFYCTHFNRRNQFFKKHEFVSNKNNISAMDRAQTIFN 278  
 85 .....QIKQQVQLIKK...VGSKVEKDLNLNEDENKKN 114  
 279 IFRFNIRKKLKDVKIEKIAVYMLEVKDFNFNYYLTKSCPLPENWRERKQ 328  
 115 GLSEQQVKEEQLRTITEEQVKYQNLVFNMNYQLDLNESGGHRRHRETDY 164  
 329 KIENLINKTREEKSKYYEELFSYTTDNKCVTQFINE. FFYNILPKDFLTG 377  
 165 DTEKWFEISHDQK.....NYVSIYANQKTSYCWWLKDYFNK 200  
 378 RNRKNFQKKVKKYVELNKHELIHKNLLKINTREISWMQVETSAHFYY 427  
 201 NNYDHLNVSINRLE..TEAEFYAFDDFSQTIKLTNNNSYQTVNID..... 242  
 428 FDHENIYVLWKLLRWI..FEDLVVSLIRCFYYVTEQQKSYSKTYYYRKNI 475  
 243 VNFDDNNLCILALLRFLLSLERFNILNIRSSY..TRNQYNFEKIGELLETI 290  
 476 WDViMKMSIADLKETLAEVQEKEVEEWKSLGFAPGKRLIPKKTTFRP 525  
 291 FAVVFSHR.....HLQGIHLQVPCAEFQYLVNSSSQISVKDSQLQ 330  
 526 IMTFNKKIVNSDRKTTKLTNTKLLNSHMLKTLKNRMFKDPFGFAVFNY 575  
 331 VYSFSTDLKLV..TNKVQDYFKFLQEFPRLTHVSQQAIPVSATNAVENL 378

FIG. 14A



16/78

576	DDVMKKYEEFVCKWKQVGQPKLF. . . . .	FATMDIEKCYDS..VNREK	615
379	NVLLKKVKH ANNLNLVSIPTQFNDFYFVNQHLKLEFGLEPNILTQKQK		426
516	LSTFL. . . . . KTTKLLSSDFWIMTAQILKRKNM. . . . .	VIDSKNFRKEMK	657
427	LENLLSIKQSKNLKFLRLNFYTYVAQETSRKQILKQATTIKNLKNNKNQ		476
558	DYFRQKFQKIALEGGQYPTLFSVLEN. . . . .	EQNDLNAKTLIVEAKQRNYFK	705
477	EETPETKDETPSESTSGMKFFDHLSELTELEDFSVN. . . . .	LQATQEYIY..	520
706	KDNLLQPVINICQYNYINFNGKFYKQTKGIPQGLCVSSILSSFYATLEE		755
521	.DSLHKLLIRSTNLKKFKLSSYKEMEKSMDTFLDLKNI. . . . .	YETLNN	564
756	SSLGFLRDESMNPENPNVNLLMRLTDDYLLITTQENNAVLFIEKLINVR		305
565	....LKRCSVNISNPHGNISYELTN. . . . .	KDSTFYKFKLTNQE	500
806	ENGFKFNMKKLQTSFPLSPSKFAKYGMDSVEEQNIVQDYCDWIGISIDMK		855
601	LQHAKYTFK. . . QNEFQFNNVKSAKIESSSLESLEDIDS LCKSIASCKNLQ		648
856	TLALMPNINLRIEGILCTLNLNMQT..KKASMWLKK..KLKSFLMNNITH		901
649	NVNI. . . . . IASLLYPNNIQKNPFPNKPNNLFFKQFEQLKNLENVSINC		691
902	YFRKTI. . . TTEDFANKTLNKLFISSGGYKYMCAKEYKDHFKNLAMSSM		948
692	I LDQHILNSISEFLEKNKKIKAFILKRYLLQYYLDYTKLFKTLQQLPEL		741
949	IDLEVSKIIYSVT. . . . . RAFFKYLVCNIKDT..IFGEEHY		982
742	NQVYINQQLEELTVSEVHKQVWENHKQKAFYEPLCEFIKESSTQLQLIDF		791
983	PDFFLS TLKHFIEIFSTKKY IFNRVCMLKAKEAKLKSDQCQSLIQ		1028
792	DQNTVSDDS IKKILESISES SKYHHYRLNPSQSSSLIKSEN EEEIQELLK		840

FIG. 14B

4	DIDLDIENLLPNTFNKYSSSCSDKGCKTLKSGSKSPSLTIPK. . . . .	47	
617	NVKSAKIESSSLESLEDIDS LCKSIASCKNLQNVNIIASLLYPNNIQKNP	666	
48	....LQKOLEFYFSDANLYNDSFLRKVLKSGEQRVE. . . . .	IETLLM	86
667	FNKPNNLFFKQFEQLKNLENVSINCILDQHILNSISEFLEKNKKIKAFIL		716

FIG. 15



MAY 03 2004

17/78

1	MEMDIDLDDIENL.....LPNTFNKYSSSCSDKGCKTLKGSKPS...	42
491	: .. .  ..... .. :  ..... :  ..	
	IELAIKIAVNKNLDEIKGHTAIFSDVSGSMSTSMSGAKKYGSVRTCLEC	540
43	.LTIPKLQKQ.....LEFYFS DANLYN DFLRK LVLKSGEQRV E IETLL	85
541	:   :   :   :   :   :   :   :   :   :   :   :   :	
	ALVLGLMVKQRCEKSSFYIFSSPSSQCNKCYL.EVDLPGDELRPSPMQKLL	589

FIG. 16

telomerase p43	LQKQLEFVFS	DAN	LYN	DSFLRK	LVLKSGE	QQRV	MEIETLLM				
human La	ICHQ	UEYY	FGD	FNLP	RDKFL	KEQI	.KLDEG	GWMP	LEIMIK		
Xenopus LaA	ICEQ	I	YY	FGD	HNL	RDKFL	KQQI	.LLDDG	GWMP	LEIMIK	
Drosophila La	ILRQ	VEYY	FGD	AN	LNR	DKFL	REQ	IGK	NEDG	GWVPLS	VLT
S. c. Lhp1p	CLKQ	VEFY	FSE	FN	FPY	DRFL	RTTA	EK	.NDG	GWMP	ISTIAT

FIG. 18

FIG. 19



18/78

	Motif A	Motif B	Motif C	Motif D	Motif E
Consensus telomerase p123	h--hDh--h--h	h---+QG---SP	h--YhDDhhh	h-h---X	h-hLgh-h
Dong (LINE)	GOPKLFFATMDIEKCYDSVNREKLSTFLKTTKLL-100-KFVKQTKGIPQGLCVSSILSSFFYATLEESSLGFLL	KNRNLHCTTDDYKKAFFDSIPHSWLIQVLEIYKIN-			
a1 S.c. (groupII)	FGGSNNWREVDLKKCFDTSHDLIIKEKRYISD-	28-ROJAIKKGIIYQGDSLSPWFCLALNPFLSHOLHNDR			
HIV-RT	LKKKKSVTTVLDVGDAYSFVPLDFFRKYTAFTIP-	26-HVPVGPVCVGQGAPTPALCNAVLLRDRLLAGLA			
L8543.12	VIPELTYFMKFDVKSCYDSIPRMECMRILKDALKN-	7-GIRYQYNVLPQGWKGSPAIQFQSSMTKILEPFRQN			
		68-KC1TREDGLFQGSSLSAPIVDLVYDILLEFYSEFK			
Consensus telomerase p123					
Dong (LINE)	-14-LMRLTDDYLITQENN-0-AVLFIEKLNNSRENGKFNMKKLQT-23-QDYCDWIGISI				
a1 S.c. (groupII)	-16-HLIYMMDDIKLYAKNDKE-0-MKKLIDTTTIFSNDISMQFGLDKCKT-25-KCLYKYLGFQO				
HIV-RT	-55-YVRYADDLILGIGLSKN-2-KIIKRDLNFLNS.LGLLTINEEKTLI- 4-ETPARFLGYNI				
L8543.12	- 4-1YQYMDDLYVGSHLEIG-1-HRTKIEELRQHILRWGLTTPDKXHQK- 0-EPPFLWMGYEL				
	- 8-ILKIAADDFLIISTDQQQ.....VINIKKIAMGGFQKYNAKANR-41-IRSKSSKGIFR				

FIG. 17



19/78

MEIENNQAQQPKAEKLWWELELEMQENQNDIQVRVKIDDPKQYL  
 VNVTAACLLQEGSYYQDKDERRYIITKALLEVAESDPEFICQLA  
 VYIRNELYIRTTNYIVAFCVVHKNTQPFIEKYFNKAVLLPNDL  
 LEVCEFAQVLYIFDATEFKNLYLDRILSODIRKELTFRKCLQRC  
 VRSKFSEFNEYQLGKYCTESQRKKTMFRLSVTNKQKWDQTKKK  
 RKENLLTQLQAIKESEDKSKRRETGDIMNVEDAIAKALKPAVMKKI  
 AKRQNAMKKHMKAPKIPNSTLESKYLTFKDLIKFCHISEPKERV  
 YKILGKKYPKTEEEYKAAGFDSASAPFNPELAGKRMKIEISKTW  
 ENELSAKGNTAEVWDNLISSNQLPYMAMLRNLSNILKAGVSDTT  
 HSIVINKICEPKAVENSKMFPLQFFSAIEAVNEAVTKGFKAKKR  
 ENMNLKGQIEAVKEVVEKTDEEKDMELEQTEEGEFVKVNEGIG  
 KQYINSIELAIIKAVENKNLDEIKGHTAIFSDVSGSMSTSMSGGA  
 KKYGSVRTCLECALVLGLMVKQRCEKSSYIIFSSPSSQCNKCYL  
 EVDLPGDELRPSMQKLLQEKGKLGGGTDFPYECIDEWTKNKTHV  
 DNIVILSDMMIAEGYSDINVRGSSIVNSIKKYKDEVNPNIKIFA  
 VDLEGYGYKCLNLDGFNENNYIKIFGMSDSILKFISAKQGGANM  
 VEVIKNFALQKIGQK

FIG. 20

MSRRNQKKPQAPIGNETNLDFVLQNLQVYKQSIEHYKTQQQQIK  
 EEDLKLLKFKNQDQDGNSGNDDEENNSNKQQELLRRVNQIKQ  
 QVQLIKKVGSKVEKDLNLNEDENKNGLSEQQVKEEQLRTITEE  
 QVKYQNLVFNMDYQLDLNESGGHRRRRETDYDTEKWFELSHDQ  
 KNYVSIYANQKTSYCWNLKDYFNKNNYDHLNVSIINRLETEAEFY  
 AFDDFSQTIKLTNNSYQTVNIDVNFDNNLICILALLRFLLSLERF  
 NILNIRSYTTRNQYFEKIGELLETIFAVVFSHRHLQGIIHLQVP  
 CEAQYLVNSSSQISVKDSQLQVYSFSTDKLVLDTNKVQDYFKF  
 LQEFPRLTHVSQQAIPVSAVNAVENLNVLKKVKHANLNLVSI  
 TQFNFDFYFVNQQLKLEFGLEPNILTQKLENLLSIKQSKNL  
 KFLRLNFYTYVAQETSRSRKQILQKATTIKNLKNKNQEEETPETKD  
 ETPSESTSGMKFFDHLSELTELEDFSVNLQATQEIYDSLHKLLI  
 RSTNLKKFKLSYKEMEKSMDTFLDLKNIYETLNILKRCSVNI  
 SNPHGNISYELTNKDDSTFYKFKLTLNQELQHAKYTFKQNEFQFN  
 NVKSAKIESSSLESLEDIDSCKSIASCKNLQNVNIIASLLYPN  
 NIQKNPFPNKPNNLFFKQFEQLKNLENVSINCILDQHILNSISEF  
 LEKNKKIKAFLKRYYLLQYLDYTKLFKTLQQLPELNQVYINQ  
 QLEELTVSEVHKQVWENHKQKAFYEPLECEFIKESSQTLQLIDFD  
 QNTVSDDSIKKILESISESKYHHLRLNPSQSSSLIKSENEEIQ  
 ELLKACDEKGVLVKAYYKFPLCLPTGTYYDYNDRW

FIG. 22

MKILFEFIQDKLDIDLQTNSTYKENLKGHFNGLDEILTCFAL  
 PNSRKIALPCLPGDLSHKAVIDHCIYLLTGELYNNVLTFGYKI  
 ARNEDVNNSLFCHSANVNVTLLKGAAWKMFSVGVYAFVDLI  
 NYTVIQFNGQFFTQIVGNRCNEPHLPPKWBQRSSSSATAAQIK  
 QLTEPVTKNQFLHKLNISSSSFPYSKILPSSSIKKLTDLREA  
 IFPTNVLKIPQRLKVRINLTLQKLLKRHKRLNYVSIILNSICPPL  
 EGTVLDLSHLSRQSPKERVLKFTIIVILQKLLPQEMFGSKKNKGK  
 IIKNLNLNLLSLPLNGYLPFDSSLKKLRLKDFRWLFIISDIWFTKH  
 NFENLNQLAICFISWLFRQLIPKIIQTFFYCTEISSTVTIVYFR  
 HDTWNKLITPFIIVEYFKTYLVEVNVCRNHSYTLNSNFHNSKMR  
 IIPKKSNNFRIIAIPCRGADEEEFTIYKENHKNAIQPTQKILEY  
 LRNKRPTSFCKIYSPQTQIADRIKEFKQRLKKFNNVLPELYFMK  
 FDVKSCYDSIPRMECMRILKDALKNENGFFVRSQYFFNTNTGVL  
 KLFNVVNASRVPKPYELYIDMVRTVHLSNQDVINVVEMEIFKTA  
 LWVEDKCYIREDGLFQGSSLASAPIVDLVYDDLEFYSEFKASPS  
 QDTLILKLADDFLIISTDQQQVINIKKIAMGGFQKYNAKANRDK  
 ILAVSSQSDDDTIVQFCAMHIFVKELEVWKSSTMNNFHIRSKS  
 SKGIFRSLIALFNTRISYKTIDTDLNSTNTVLMQIDHVVKNISE  
 CYKSAFKDLSINVTNQMFHSFLQRIIEMTVSGCPITKCDPLIE  
 YEVRFTILNGFLESLSNTSKFKDNIIILRKEIQHLQAYIYIYI  
 HIVN

FIG. 23



20/78

1	tcaatactat	taattataaa	ataaaaaaaa	gcaaaactaca	aagaaaaatgt	caaggcgtaa
61	ctaaaaaaag	ccataggctc	ctataggcaa	tgaacaaat	cttgatttg	tattacaaaa
121	tctagaagtt	tacaaaagcc	agattgagca	ttataagacc	tagtagtaat	agatcaaaga
181	ggaggatctc	aagctttaa	agttcaaaaa	ttaagattag	gatgaaact	ctgcaacgaa
241	tgtatgtat	gaagaaaaca	actcaaataa	ataataagaa	tattattaagga	gagtcattt
301	gattaagttag	caagtttaat	tgataaaaaa	agttgggtct	aaggtagaga	aagattgaa
361	tttgaacgaa	gatgaaaaca	aaaagaatgg	actttctgaa	tagcaagtg	aagaagagta
421	attaagaacg	attactgaag	aataggtaa	gtattaaaat	tttagtattt	acatggacta
481	ccagtttagat	ttaaatgaga	gtggtgccca	tagaagacac	agaagagaaa	cagattatga
541	tactgaaaaa	tggttgaaa	tatctcatga	ccaaaaaaat	tatgtatcaa	tttagccaa
601	ctaaaagaca	tcatatttt	ggtggcttaa	agatttattt	aataaaaaca	attatgatca
661	tcttaatgtt	agcattaaca	gactagaaaac	tgaagccaa	ttctatgcct	ttgatgatt
721	ttcacaaaaa	atcaaactta	ctaataattt	ttactagact	gttaacatag	acgttaattt
781	tgataataat	ctctgtatac	tcgcattgt	tagattttt	ttatcactag	aaagattcaa
841	tatTTGAAT	ataagatctt	cttataacaag	aaattaatat	aattttgaga	aaatttggta
901	gtctacttggaa	actatcttcg	cagttgtctt	ttctcatcgc	cacttacaag	gcattcattt
961	acaagttcc	tgcgaagcgt	tctaataattt	agtttaactcc	tcatcataaa	ttagcgttaa
1021	agatagctaa	ttataagttat	actctttctc	tacagactta	aaattagttg	acactaaca
1081	agtccaaatg	tatTTTAAGT	tcttataaga	attccctctg	ttgtactatg	taagcttagt
1141	ggctatccca	gttagtgcta	ctaacgctgt	agagaacctc	aatgtttac	ttaaaaagggt
1201	caagcatgt	aatcttaatt	tagtttctat	cccttacctaa	ttcaattttt	atttctactt
1261	tgttaatttta	taacatttga	aatttagagtt	tggatttagaa	ccaaatattt	tgacaaaaca
1321	aaagcttggaa	aatctacttt	tgagttataaa	ataatcaaaa	aatcttaat	tttaagatt
1381	aaacttttac	acctacgtt	cttaagaaaac	ctccagaaaa	cagatattaa	aacaagctac
1441	aacaatcaaa	aatctcaaaa	acaataaaaa	tcaagaagaa	actctctaaa	ctaaagatga
1501	aactccaaagc	gaaaggcacaa	gtggatgaa	atTTTTGAT	catctttctg	aattaacgca
1561	gcttgaagat	ttcagcgta	acttgcataa	tacccaagaa	atttatgata	gcttcaccaa
1621	acttttgatt	agatcaacaa	atTTAAAGAA	gttcaaatta	agttacaat	atggaaatgg
1681	aaagagtaaa	atggatacat	tcatagatct	taagaatatt	tatgaaacct	taaacaatct
1741	taaaagatgc	tctgttataa	tatcaatattc	tcatggaaac	atttctttagt	aactgacaaa
1801	taaagattct	acttttttta	aattttaagct	gacccttaac	taagaattat	aacacgctaa
1861	gtatactttt	aagtagaacg	aTTTTTAATT	taataacggt	aaaagtgc	aaattgaatc
1921	ttcctcattt	gaaagcttag	agatattttg	tagttttgc	aatatcttgg	cttcttggta
1981	aaattttacaa	aatgttataa	ttatcgccag	tttgccttat	cccaacaata	tttagaaaaaa
2041	tcctttcaat	aagcccaatc	ttcttattttt	caagcaattt	gaataatgt	aaaatttggaa
2101	aatgttatct	atcaactgt	ttcttgcata	gcataatactt	aattcttattt	cagaatttctt
2161	aaaaaaagaat	aaaaaaataa	aagcattcat	tttggaaaaga	tattatttt	tacaatattt
2221	tcttgattat	actaaattat	ttaaaaacact	tcaatagtt	cctgaattaa	attaagttt
2281	cattaattag	caatttagaa	atttgactgt	gagtgaaagta	cataagtaag	tatggggaaaa
2341	ccacaagcaa	aaagctttct	atgaaaccatt	atgtgagtt	atcaaagaat	catctttaac
2401	cctttagctt	atagattttt	acccaaaacac	tgtaaagtgtat	gactcttattt	aaaagattttt
2461	agaatctata	tctgagttct	agtatcatca	ttatttgaga	ttgaacccta	gttaatctag
2521	cagtttaattt	aaatctgaaa	acgaagaaaat	ttaagaactt	ctcaaagctt	gcgcacgaaaa
2581	aggtgttttta	gtaaaagcat	actataaaatt	ccctctatgt	ttaccaactg	gtacttattt
2641	cgattacaat	tcatagat	ggtgattaaat	taatatttag	tttaaataaaa	tattaaatat
2701	tgaatatttc	tttgcttattt	atttgaataaa	tacataacaat	agtcattttt	agtgttttga
2761	atataatttttta	gttattttat	tcattattttt	aagttaaataaa	ttatTTTTCA	atcatTTTT
2821	aaaaaaatcg					

FIG. 21



21/78

Oxytricha  
Eupletes

LCVSYILSSFYANLEENALQFLRKE\$MDPEK\$ETNLLMRLT  
LCVSSILSSFYATLEESSLGFLRDESMNPENPNVNLLMRLT

FIG. 24

ATTTATACTCATGAAAATCTTATTGAGTCATTCAAGACAAGCTTGACATTGATCTACA  
GACCAACAGTACTTACAAAGAAAATTAAATGTGGTCATTCAATGGCCTCGATGAAAT  
TCTAACTACGTGTTCCGACTACCAAATTCAAGAAAAATAGCATTACCATGCCTTCTGG  
TGACTTAAGCCACAAAGCAGTCATTGATCAGTCATCATTACCTGTTGACGGGCGAATT  
ATACAACAACGTACTAACATTGGCTATAAAATAGCTAGAAATGAAGATGTCAACAATAG  
TCTTTTTGCCATTCTGCAAATGTTAACGTTACGTTACTGAAAGGCCTGCTTGGAAAAT  
GTTCCACAGTTGGTCGGTACATACGCATTGTTGATTATTGATCAATTATACAGTAAT  
TCAATTAAATGGGCAGTTTCACTCAAATCGTGGGTAACAGATGTAACGAACCTCATCT  
GCCGCCAAATGGGTCAACGATCATCCTCATCATCCGCAACTGCTGCGCAAATCAAACA  
ACTTACAGAACCGAGTGACAATAAACAAATTCTTACACAAGCTCAATATAAATTCCCTTC  
TTTTTTCCATTAGCAAGATCCTTCCTCATCATCTATCAAATCAGACTAAGCTACTGACTT  
GAGAGAAGCTATTTCACAAATTGGTTAAAATTCCCTCAGAGACTAAAGGTACGAAT  
TAATTGACGCTGCAAAGCTATTAAAGAGACATAAGCGTTGAATTACGTTCTATT  
GAATAGTATTGCCCCACATTGGAAGGGACCGTATTGGACTTGTGCAATTGAGTAGGCA  
ATCACCAAAGGAACGAGTCTTGAATTATCATTTGTTATTACAGAAGTTATTACCCCA  
AGAAATGTTGGCTCAAAGAAAATAAGAAAATTATCAAGAATCTAAATCTTTATT  
AAGTTACCCCTAAATGGCTATTACCATTTGATAGTTGTTGAAAAGTTAAGATTAAA  
GGATTTCGGTGGTTGTCATTCTGATATTGGTTACCAAGCACAATTGAAAACCTT  
GAATCAATTGGCGATTGTTCTTCTGGCTATTAGACAACATTCCAAAATTAT  
ACAGACTTTTTACTGCACCGAAATATCTCTACAGTGACAATTGTTACTTAGACA  
TGATACTTGAATAAACCTTACACCCCTTATCGTAGAATTAAAGACGTACTTAGT  
CGAAAACAACGTATGTAGAAACATAATAGTTACACGTTGTCACATTCAATCATAGCAA  
AATGAGGATTATACAAAAAAAGTAATAATGAGTCAGGATTATTGCCATCCCATGCG  
AGGGGCAGACGAAGAAGAATTACAATTATAAGGAGAATCACAAAATGCTATCCAGCC  
CACTAAAAAATTAGAACACCTAACAGAACAAAAGGCCACTAGTTACTAAATATA  
TTCCTCAACGCAAATAGCTGACCGTATCAAAGAACATTAAAGCAGAGACTTTAAAGAAATT  
TAATAATGCTTACCAAGAGCTTATTGATGTCACATTGCTATGATT  
CATACCAAGGATGGAATTGATGAGGAACTCAAGGATGCGCTAAAAAATGAAAATGGGTT  
TTCGTTAGATCTCAATTCTCAATACCAATACAGGTGTATTGAAGTTATTAAATGTT  
TGTTAACGCTAGCAGAGTACCAAAACCTTATGAGCTACATAGATAATTGAGGACGGT  
TCATTATCAAATCAGGATGTTAAACGTTGAGAGATGGAATTTTAAACAGCTTT  
GTGGGTTGAAGATAAGTGTACATTAGAGAAGATGGCTTTTCTAGGGCTCTAGTTATC  
TGCTCCGATCGTTGATTGGTGTATGACGATCTCTGGAGTTTATAGCGAGTTAAAGC  
CAGTCCTAGCCAGGACACATTAATTAAAAGCTGGCTGACGATTCTTATAATATCAAC  
AGACCAACAGCAAGTGTACATATCAAAGCTTGCCATGGCGGATTCAAAAATATAA  
TGCAGAACAGCAATAGAGACAAAATTAGCCGTAAGCTCCCAATCAGATGATGACGGT  
TATTCAATTGTCGAATGCACATATTGTTAAAGAATTGGAAGTTGAAACATCAAG  
CACAATGAATAATTCCATATCCGTTGAAATCTAGTAAAGGGATATTGCAAGTTAAT  
AGCGCTGTTAACACTAGAAATCTCTATAAAACAATTGACACAAATTAAATTCAACAAA  
CACCCTCTCATGCAAATTGATCATGTTGAAAGAACATTGGAATGTTATAAATCTGC  
TTTAAGGATCTATCAATTAAATGTTACGCAAATATGCAATTCAATTGCTTACAACG  
CATCATTGAAATGACAGTCAGCGGTTGTCACAGAAATGTGATCCTTAATCGAGTA  
TGAGGTACGATTCAACATATTGAATGGATTGTTGAAAGCCTATCTCAAACACATCAA  
ATTAAAGATAATATCATTCTTGAGAAAGGAATTCAACACTGCAAGC

FIG. 26



22/78

AKFLHWLMSVYVVELLRSFFEFYVTEFTFQKNR  
LSEIEWLVLGKRSNAKMCCLSDFEKRKQIQAEEFIYWLNSFLIPILOQSFYITESSDLRNR  
LKDFRWLFISD---IWFTKHNFENLNQNLIAICFISWLFRQLIPIKIIQTFFYCTEISSTVT-  
TREISWMQVET-SAKHFFYFDHEN-IIYWLKLLWIFEDVVSLLRCEQQKSYSK  
\*\*\* \*\*\*  
\* \*

human  
tez1  
EST2  
p123

human  
tez1  
EST2  
p123

Motif 2  
RPIVNMDYVVGARTFRREKRAERLTSRVKALF-SVLYNERYA  
RLITN-LRKRFKLKGMSNKKMVLVSTNQTLRPVASILKHLINNEESSGGIPFNLEVYMKLTLF  
RIIAIPCRGADEEEFTIYKENHKNALOPTOKILLEYLRNKRPSTSFTKLYSPTQIADRKEF  
RPIIMTFNKKLTVNSDRKTTKLTTNTKLUNSHMLKTLKN-RMFKDOPFFGAEVNYDDVMKKY  
\* \* \*

tez1  
EST2  
p123

FIG. 25

1



23/78

AKFLHWLMSVYVVELLRSFFYVTETTFQKNRLFFYRKSVWSKLQSIGIRQHLKR  
VQLRDVSEAEVRQHREARPALLTSRLRFIPKPDGLRPIVNMDYVGARTFRREK  
RAERLTSRVKALFSVLNYERA

*FIG. 27*

GCCAAGTTCCCTGCACTGGCTGATGAGTGTACGTCGAGCTGCTCAGGTC  
TTCTTTTATGTCACGGAGACCACGTTCAAAAGAACAGGCTCTTTCTACC  
GGAAGAGTGTCTGGAGCAAGTTGCAAAGCATTGGAATCAGACAGCACTTGAAG  
AGGGTGCAGCTGCGGGACGTGTCGGAAGCAGAGGTCAAGCAGCATCGGGAAAGC  
CAGGCCCGCCCTGCTGACGTCCAGACTCGCTTCATCCCAAGCCTGACGGGC  
TGCAGCCGATTGTGAACATGGACTACGTCGTTGGAGCCAGAACGTTCCCGAGA  
GAAAAGAGGGCCGAGCGTCACTCGAGGGTGAAGGACTGTTCAGCGTGCT  
CAACTACGAGCGGGCGCG

*FIG. 28*

MTEHHTPKSRILRFLENQYVYLCTLNDYVQLVLRGSPASSYSNICERLRSVDQTSFSIFLHSTVVGF  
DSKPDEGVQFSSPKCSQSELIANVVKQMFDESFERRRNLLMKGFSMNHEDFRAMHNGVQNDLVSTF  
PNYLISILESKNWLLEIIGSDAMHYLLSKGSIFEALPNDNYLQISIGIPLFKNNVFEETVSKKRKR  
TIETSITQNKSARKEVSWNSISISRFSIFYRSSYKKFKQDLYFNLHSICDRNTVHMWLQWIFPRQFG  
LINAFQVKQLHKVIPLVSQSTVVPKRLKVYPLIEQTAKLHRISLSKVYNHYCPYIDTHDDEKILS  
YSLKPNQVFAFLRSILVRFPKLIWGNQRIFEIILKDETLFLKLSRYESFSLHYLMSNIKISEIEWL  
VLGKRSNAKMCLSDFEKRQIIFAEFIYWLNSFIIPILQSFFYITESSDLRNRTVYFRKDIWKLLCR  
PFITSMKMEAFAEKINENNVRMDTQKTLPPAVIRLLPKKNTFRLITNLRKRFLIKMGSNKKMLVSTN  
QTLRPVASILKHLINNEESSGIPFNLEVYMKLLFKKDLLKHRMFGRKKYFVRIDIKSCYDRIKQDLM  
FRIVKKKLKDPEFVIRKYATIHATSDRATKNFVSEAFSYFDMVPFEKVVQLLSMKTSDTLFVDFVDY  
WTKSSEIFKMLKEHLSGHIVKIGNSQYLQKVGIPQGSILSSFLCHFYMEDLIDEYLSFTKKKGSVL  
LRVVDFLFITVNKKDAKKFLNLSLRGFEKHNFSTSLEKTVINFENSNGIINNTFFNESKKRMPFFG  
FSVNMRSLDTLLACPKIDEALFNSTSVELTKHMGKSFFYKILRSSLASFAQVFIDITHNSKFNSCCN  
IYRLGYSMCMRAQAYLKRMDIFIPQRMFITDLLNVIGRKIWKKLAELGYSRRFLSSAEVKWLFC  
LGMRDGLKPSFKYHPCFEQLIYQFQSLTDLIKPLRPVLRQLFLHRRIA

*FIG. 29*



24/78

FIG. 30A

1



FIG. 30B



26/78

EST2	pep	NVCRNHSNSY-	-----	TLNFNHSRM	RITPKKSNNNE	FRITAIPCRG	129
Euplotes	pep	KEVEEWKSL	-----	-----	RLIPPKRTT--	FRPIMTFNKK	120
Trans of tetrahymen		KIQLEENLE	KVEKLIPED	-----	RLIPKKGS--	FRPIMTFLRK	130
Consensus		K...E....	-----	-----	F...GKL	FRPIMTF.RK	100
EST2	pep	ADEEEFTIYK	ENHKNAIQPT	OKILEYTRNK	RPTSTFTKIS	PTQIADRIKE	40
Euplotes	pep	IVNSDRKTTK	LTTNTKLNS	HLMLKTLKN-	-----	-RMFK	43
Trans of tetrahymen		DKQKNIK--	--LNLNQILMDS	QLVFNKLD-	-----	-DPFGFAVFN	44
Consensus		.....K..K	LN.N..L..S	QL.L..IKN-	-----	-ML-G	44
EST2	pep	FKQRLKKFN	NVL--	-----	-FELYFMKFD	VKSCYD	157
Euplotes	pep	YD-DVMKYE	EFVCKWKQVH	CPKLUFFATMD	IEKCYD	155	
Trans of tetrahymen		NK-QISEKFA	QFIEWKWNKG	RPCLQYVTL-	-----	158	
Consensus		.K.....KKF.	.F..KWK..G	P..LYF.T.D	...CYD	186	

FIG. 31



27/78

S-1: FFY VTE TTF QKN RLF FYR KSV WSK  
 S-2: RQH LKR VQL RDV SEA EVR QHR EA  
 S-3: ART FRR EKR AER LTS RVK ALF SVL NYE

A-1: AKF LHW LMS VYV VEL LRS FFY VTE TTF Q  
 A-2: LFF YRK SVW SKL QSI GIR QHL KRV QLR DVS  
 A-3: PAL LTS RLR FIP KPD GLR PIV NMD YVV

*FIG. 32*

Poly 4

t t c  
 t a a g c c t c g  
 5' - c a g a c c a a a g g a a t t c c a t a a g g - 3'  
 Q T K G I P Q G

4 (B')

5 (c')

D D Y L L I T  
 3' - c t g c t g a t g g a g g a g a t a g t g g - 5'  
 a a a a a a a a a a  
 t t t t  
 c c  
Poly 1

*FIG. 34*

+

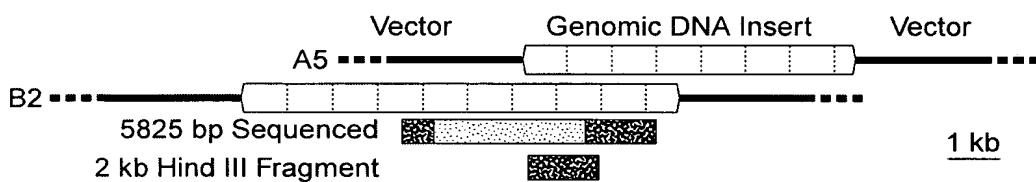


FIG. 33A

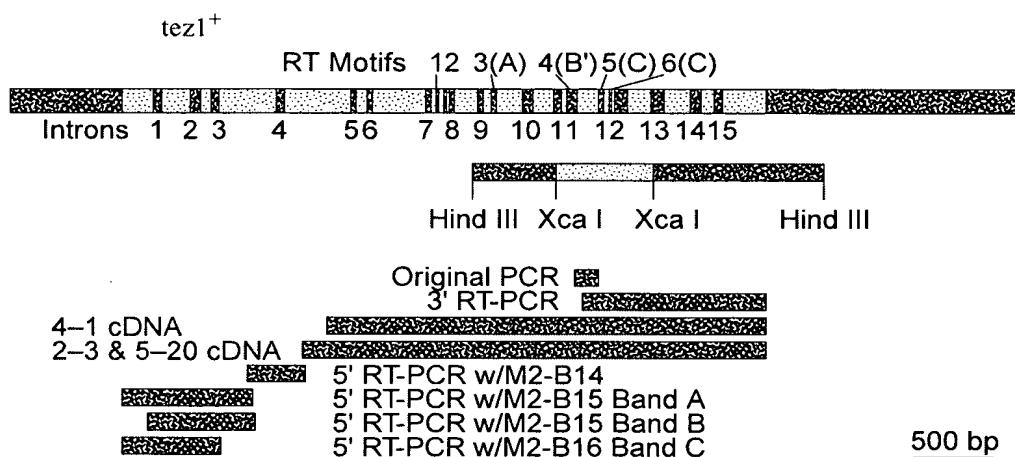


FIG. 33B



29/78

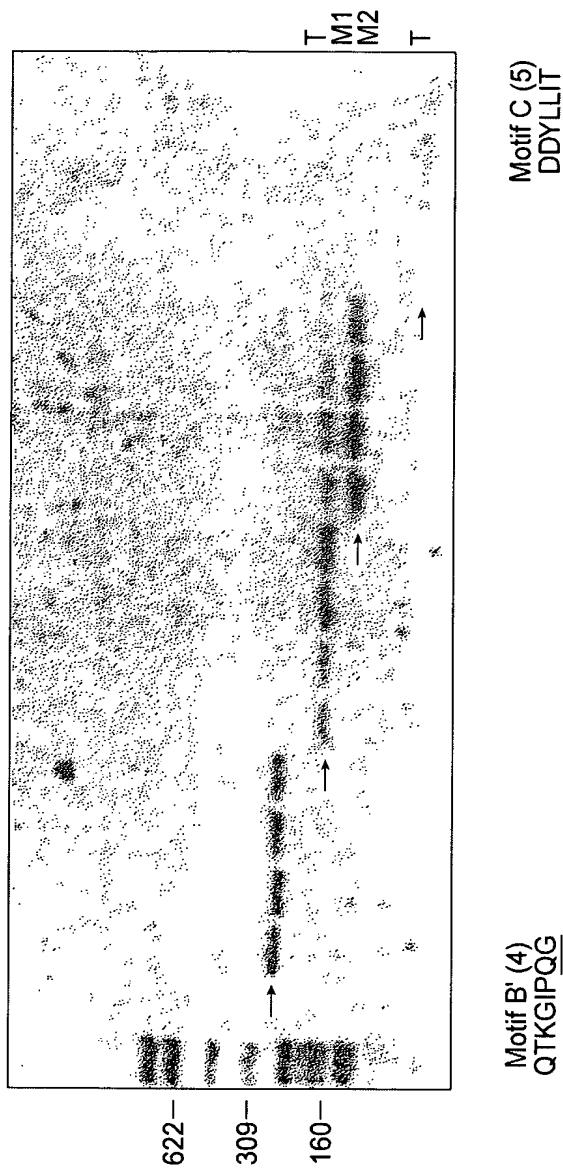


FIG. 35



30/78

Ot	LCVSYIILSSFYANLEENALQFLRKESMDPEKPKETNLLMRLT
Ea_p123	KGIPQGLCVSSSISSFYATLEESSLGFLRDESMNPENPNVNLLMRLTDDYLLIT
Sp_M2	SILSSFLCHFYMEDLIDEYLSFTKKK-----GSVLLRVV
Sc_p103	DGLFQGSSLSSAPIVDLVYDDILEFYSEFKASPS-----QDTLILKLAADDLIIIS

\* . . . . \*

Q K V G I P Q G  
caa aaa gtt ggt atc cct cag gg..... <--Actual Genomic Sequence.

Poly 4  
t t t c  
t a a g c c t c g  
cag acc aaa gga att cca taa gg ----->

ag acc aaa gga att cca tca ggc tca ATT CTG TCA TCT TTT TTG TGT CAT TTC TAT ATG  
tc tgg ttt cct taa ggt agt cgg AGT TAA GAC AGT AGA AAA AAC ACA GTA AAG ATA TAC  
K G I P S G S I L S S F L C H F Y M

F/G. 36A



31/78

GAA GAT TTG ATT GAT GAA TAC CTA TCG TTT ACG AAA AAG AAA GGA TCA GTG TTG TTA CGA  
CTT CTA AAC TAA CTA CTT ATG GAT AGC AAA TGC TTC TTT CCT AGT CAC AAC AAT GCT  
E D L I D E Y L S F T K K G S V L L R

GTA GTC gac gag tac ctc ctc atc acc  
CAT CAG ctg ctg atg gag gag tag tag tgg

V V D D Y L L I T

<---- ctg ctg atg gag gag tag tag tgg  
a a a a a a a a  
t t t t t t  
c c  
Poly 1

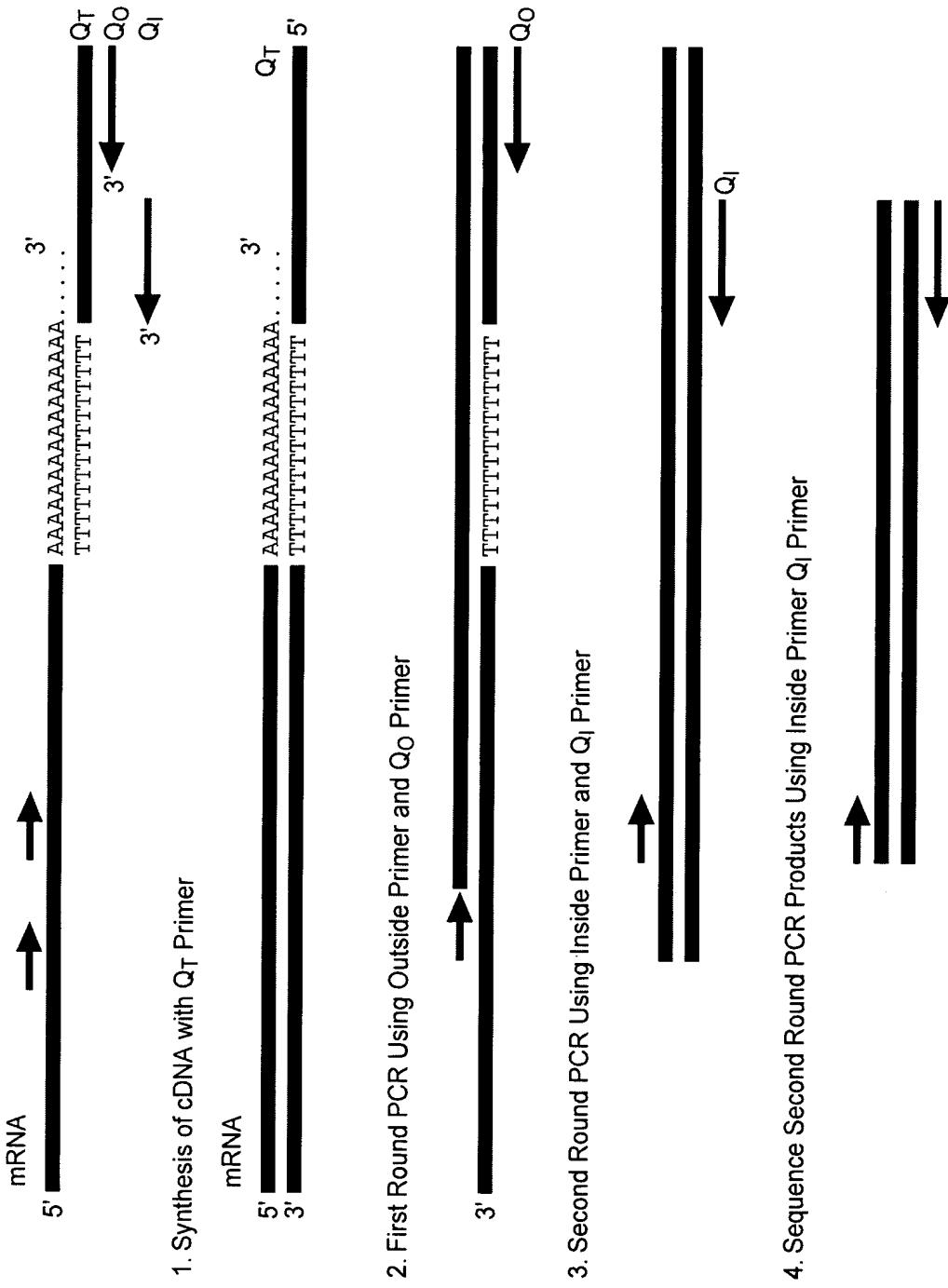
.....gac gat ttc ctc ttt ata aca.....<----Actual Genomic Sequence  
D D F L F I T

FIG. 36B

+



32/78





33/78

**A. Genomic Libraries**

Size Selected Libraries from *P. Nurse*

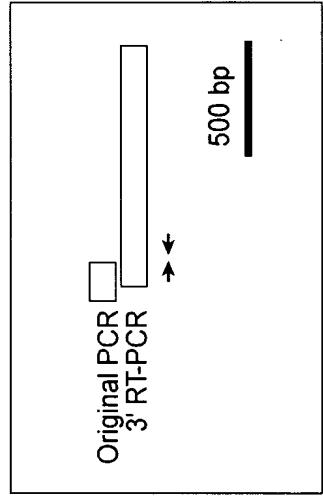
- 3~4 kb
- 5~7 kb
- 7~8 kb
- 11~12 kb

Libraries from J.A. Wise

- Sau 3a Partial Digest
- Hind III Partial Digest

cDNA Libraries

- GAD (Gal Activation Domain) Library
- REP Library from R. Alshire
- REP81ES Library (old)
- REP81ES Library (new)
- REP41ES Library



**B.**

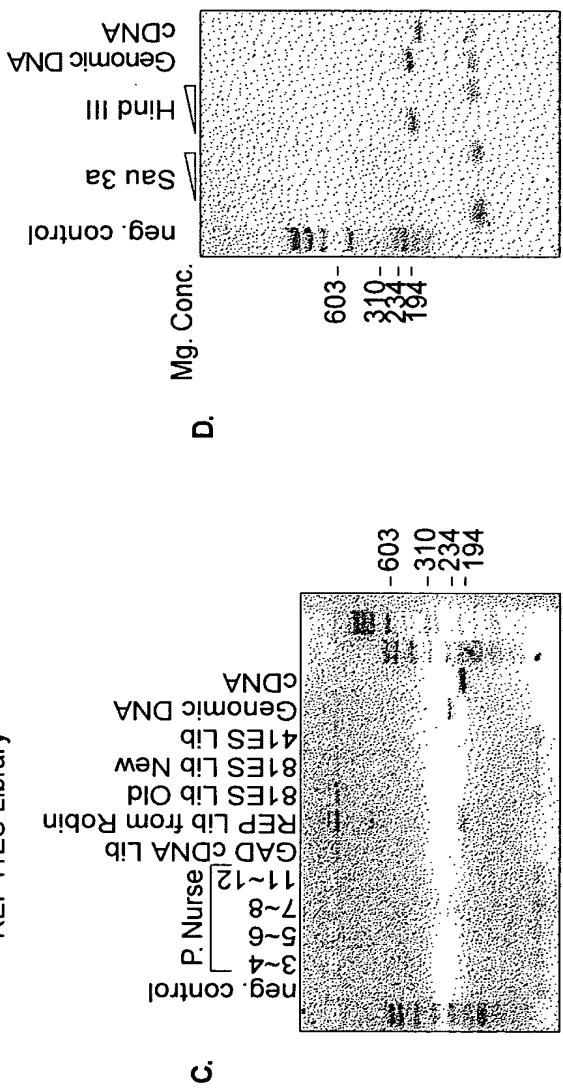


FIG. 38

+

O I P E JC13  
MAY 03 2004  
PATENT & TRADEMARK OFFICE

+

34/78

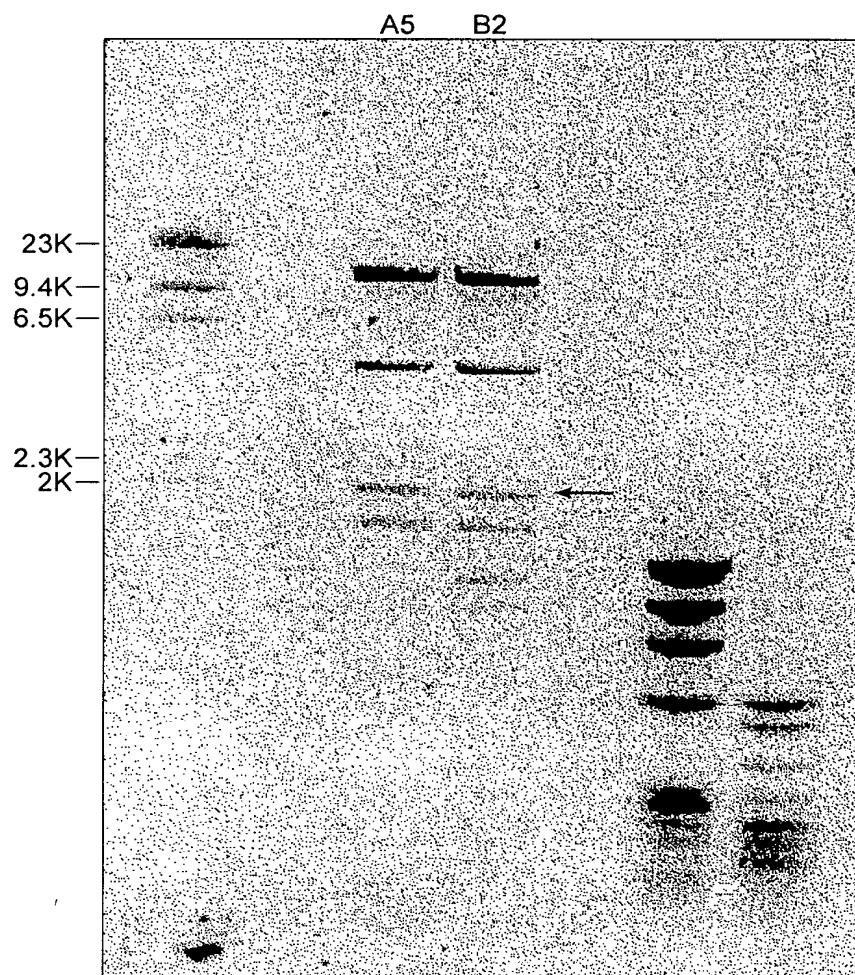
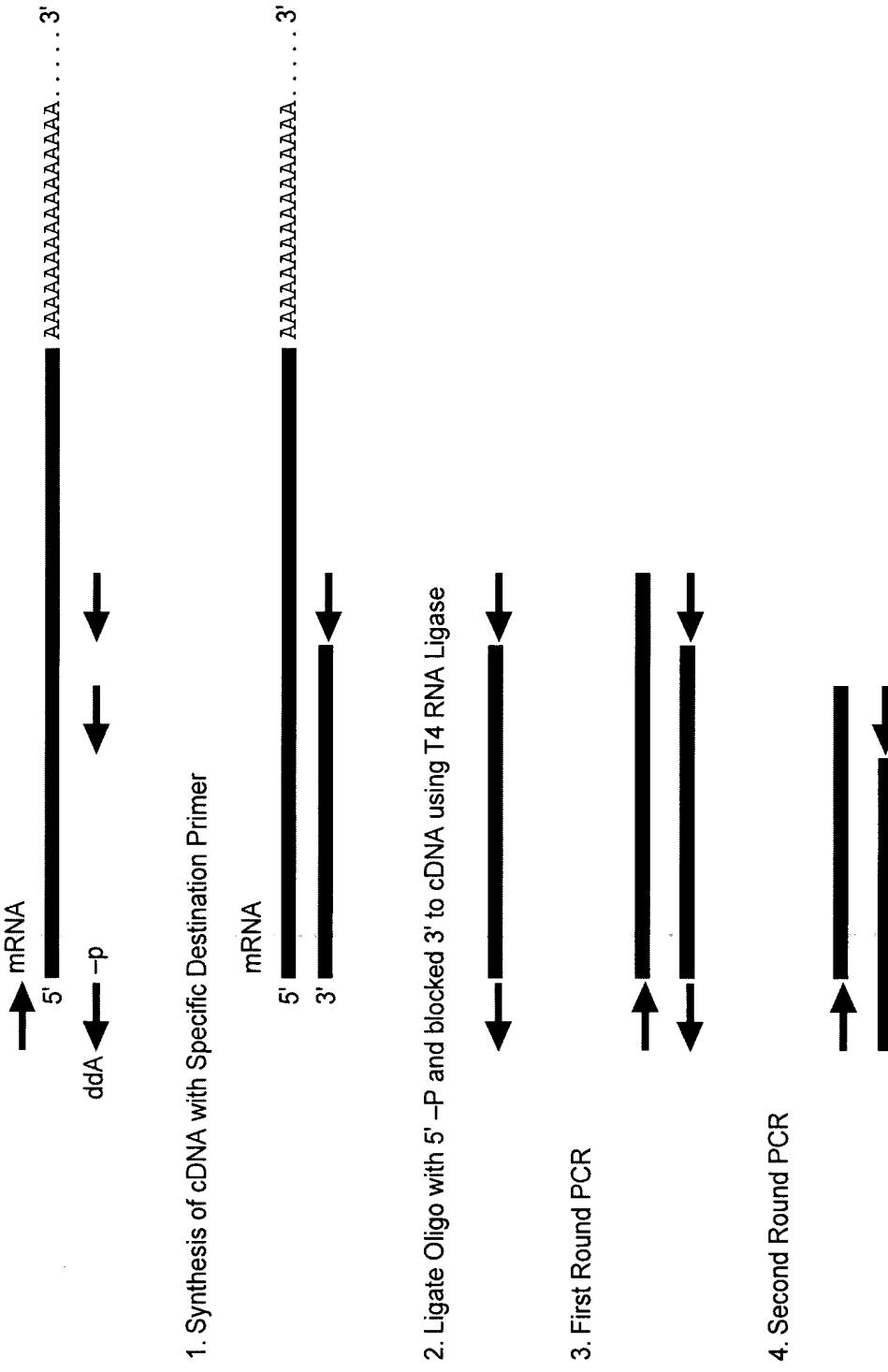


FIG. 39

+



35/78



F/G. 40

+



36/78

S.p.	Tez1p	(429).	WLYNSFIPIIPIQSQFFYITESSDLRNRTVYFRKDIW	... (35) ...	Motif 0
S.c.	Est2p	(366).	WLFRQLIPKIQTFFYCYTEISSTVT-IVYFRHDTW	... (35) ...	
E.a.	p123	(441).	WIFEDLVVSLIRCFYYTEQQSYSKYYRKNIW	... (35) ...	
* * * * *					
			Motif 1	Motif 2	
			P hh h K	hR h R	
S.p.	Tez1p	AVIRLLPKK--NTFRLITN-LRKRF	... (61) ...		
S.c.	Est2p	SKMRRIPKKSNNNEFRIIAIPCRGAD	... (62) ...		
E.a.	p123	GKLLRIPKK--TTFRPIMTENKKIV	... (61) ...		
* * * * *					
			Motif 3 (A) AF		
			h hDh GY h		
S.p.	Tez1p	KKYFVRIDIKSICYDRIKQDLMFRIVK	... (89) ...		
S.c.	Est2p	ELYFMKEDVKSCYDSIPRMECMRLK	... (75) ...		
E.a.	p123	KLFFATMDIEKCYDSVNREKLSTFLK	... (107) ...		
* * * * *					
			Motif 4 (B')		
			hPQG pp hh h		
S.p.	Tez1p	YLQKVGVIPQGSILSSFLCHFYMEDLIDEYLSF	... (6) ...		
S.c.	Est2p	YIREDGFLFGQSSILSAPIVDLVYDDLLEFYSEF	... (8) ...		
E.a.	p123	YKQTKGIPQGLCVSSILSSFYATLLESSLGF	... (14) ...		
* * * * *					
			Y Motif 5 (C)		Motif 6 (D)
			h F DDhhh	Gh h CK h	
S.p.	Tez1p	VLLRVDDFLFITVNKDAKKFLNLSLRGFEKHNFSTSLEKTVINFENS	... (205) ...		
S.c.	Est2p	LILKLADDFLIISTDQQQVINIKKLAMGGFOQYNAKANRDKILAVSSQS	... (173) ...		
E.a.	p123	LLMRLTDDYLITTOENNNAVLFIEKLINVSRENGFKFNMKKLQTSFPLS	... (209) ...		
* * * * *					

FIG. 41



37/78

A.

Sp_Tip1p	1	- - - - -	M T E H H T P K S R I L R F L E N Q Y Y L C T	24
Sc_Est2p	1	- - - - -	M E V D V D N Q A D N H G I H S A L K T C E E I K E A K T L Y S W	7
Ea_p123	1			
Sp_Tip1p	25	L N D Y V Q L V L R G S P A S S Y S N I C E R L R S D V Q T S F S	57	
Sc_Est2p	8	I Q D K L D I D L Q T N - - S T Y K - - E N L K C G H F N G L D	35	
Ea_p123	34	I Q K V I R C R N Q S Q - - S H Y K - - D L E D I K I F A Q T N	61	
Sp_Tip1p	58	I F L H S T V V G F D S K P D E G V Q F S S P K C S Q Q S E L I A N	90	
Sc_Est2p	36	E I L T T C F A L P N S R - K I A L P C L P G D L S H K A V I D H	67	
Ea_p123	62	I V A T P R D Y N E E D F K V I A R K E V F S T G L M I E L I D K	94	
Sp_Tip1p	91	Y Y K Q M F D E S F E R R R - N L L M K G F S M N H E D F R A M H	122	
Sc_Est2p	68	C I I Y L L T G E L Y N - - N V L T F G Y K I A R N E D - - -	93	
Ea_p123	95	C L V E L L S S S D V S D R Q K L Q C F G F Q L K G N Q - - -	122	
Sp_Tip1p	123	V N G V Q N D L V S T F P N Y L I S I L E S K N W Q L L E I I G	155	
Sc_Est2p	94	- - V N N S L F C H S A N V N V T L L K G A A W K M F H S L V G	123	
Ea_p123	123	- - L A K T H L L T A L S T Q K Q Y F F Q D E W N Q V R A M I G	152	
Sp_Tip1p	156	S D A M H Y L L S K G S I F E A L P N D N Y L Q I S G I P L F K N	188	
Sc_Est2p	124	T Y A F V D L L I N Y T V I Q F N - G Q F F T Q I V G N R C N E P	155	
Ea_p123	153	N E L F R H L Y T K Y L I F Q R T S E G T L V Q F C G N N V F D H	185	
Sp_Tip1p	189	N V F E E T V S K K R K R T I E T S I T Q N - - - K S A R K E V S	218	
Sc_Est2p	156	H L P P K W V Q - R S S S S A T A A Q I - - - K Q L T E P V T	183	
Ea_p123	186	L K V N D K F D K - K Q K G G A A D M N E P R C C S T C K Y N V K	217	

FIG. 42A

+



4

Sp_Tip1p	219	WNSISISRFSSIFYRSSYKKFKQDLYFNLHSICD	251
Sc_Est2p	184	NEK-N	200
Ea_p123	218	NEK-DHFLNNINVPNWNMKSRTFYCTHFN	248
Sp_Tip1p	252	RNTVHMWLQWIFPRQFGLINAFAQVKKQLHKVIP	284
Sc_Est2p	201	YSKILPSSS-S	223
Ea_p123	249	NNQFFKKHEFVSNKNNISAMDRAQT	275
Sp_Tip1p	285	YS-QSSTVVPKRLLKVKYPLIEQTAKRLLHRI	313
Sc_Est2p	224	TVKIPQRLLKVRINLTLQKLLKRLHKRN	252
Ea_p123	276	FTNIFRFNRIRKKDKVIEKIAVYMLEKVKDFN	308
Sp_Tip1p	314	LSKVYNHCPYID-THDDDEKILSYSSLKPNQ-	342
Sc_Est2p	253	YVSILNSICPPLEGTVLDDLSSRQSPKER-	282
Ea_p123	309	FNYYLTKSCPLPENWRERKQKIENLINKTRE	341
Sp_Tip1p	343	- - - - -	359
Sc_Est2p	283	- - - - -	299
Ea_p123	342	SKYYEELFSYTTDNKCVTQFINEFFYNI	374
Sp_Tip1p	360	WGNGQRIFFEILKDLTFLKLSRYESFSSLHYLMS	392
Sc_Est2p	300	FGSKKNGKIKNLNLQKILVILQKILLPQEM	332
Ea_p123	375	LTG-RNRKNFQKKVKKYVELNKHNLNLLE	406
Sp_Tip1p	393	NIKISEIIEWLVLGKRSNAKMCLSDFEKRKQI	425
Sc_Est2p	333	KLRLKDFRWLFIS-DIWFTKHNFENLNQLA	362
Ea_p123	407	KINTREISWMQVETSAKHFFYVDFHEN-IVVLW	437

FIG. 42B



A.

Sp_Tip1p	426	EF I YWL YNS E I I PI L Q SFFF Y I TE S S D L R N R T V Y	458
Sc_Est2p	363	CF I SWL F RQL I PK I - Q T F F Y C T E I S S T V T - I V Y	394
Ea_p123	438	K L L R W I F E D L V V S L I R C F F Y V T E Q Q K S Y S K T Y Y	470
Sp_Tip1p	459	FR K D I W K L L C R P F I T S M K M E A F E K I N E N N V R M D 491	
Sc_Est2p	395	FR H D T W N K L I T P F I V E Y F K T Y L V E N N V C R N H N S 427	
Ea_p123	471	Y R K N I W D V I M K M S I A D L K K E T L A E V Q E K E V E E W 503	
Sp_Tip1p	492	T Q K T T L P P A V I R L L P K K - N T F R L I T N L R K R F L 522	
Sc_Est2p	428	Y T L S N F N H S K M R I I P K K S N N E F R I I A I P C R G A D 460	
Ea_p123	504	KK S L G F A P G K L R L I P K K - - T T F R P I M T F N K K I V 534	
Sp_Tip1p	523	I K M G S N K K M L V S T N Q T L R P V A S I L K H L I N E - - - 552	
Sc_Est2p	461	EEE - - F T I Y K E N H K N A I Q P T Q K I L E Y L R N K R P T 491	
Ea_p123	535	N S D - - R K T T K L T T N T K L L N S H L M L K T L K N R - M F 564	
Sp_Tip1p	553	E S S G I P F N L E V Y M K L L T F K K D L L K H R M F G R - K K 584	
Sc_Est2p	492	S F T K I Y S P T Q I A D R I K E F K Q R L L K K F N N V L P E L 524	
Ea_p123	565	K D P F G F A V F N Y D D V M K K Y E E F V C K W K Q V G Q P K L 597	
Sp_Tip1p	585	Y F V R I D I K S C Y D R I K Q D L M F R I V K K K L K D P E - F 616	
Sc_Est2p	525	Y F M K F D Y K S C Y D S I P R M E C M R I L K D A L K N E N G F 557	
Ea_p123	598	F E A T M D I E K C Y D S V N R E K L S T F L K T K L L S S D F 630	
Sp_Tip1p	617	V I R K Y A T I H A T S D R A T K N - - - - - 634	
Sc_Est2p	558	F V R S Q Y F F N T N T G - - - - - 570	
Ea_p123	631	W I M T A Q I L K R K N N I V I D S K N F R K K E M K D Y F R Q K 663	

FIG. 42C



40/78

A.

Sp_Tip1p	635	F V S E A F S Y F D M V P F E K V V Q L L S - - M K T S D T L F V	665
Sc_Est2p	571	· · · · · V L K L F N V V N A S R - - V P K P Y E L Y -	591
Ea_p123	664	F Q K I A L E G G Q Y P T L F S V L E N E Q N D L N A K K T L I V	696
Sp_Tip1p	666	D F Y D Y W T K S S S E I F K M L K E H L S G H I V K I G N S Q Y	698
Sc_Est2p	592	D N Y R T V H L S N Q D V I N V V E M E I F K T A L W V E D K C Y	624
Ea_p123	697	E A K Q R N Y F K K D N L L Q P V I N I C Q Y N Y I N F N G K F Y	729
Sp_Tip1p	699	L Q K V G I P Q G S I L S S F L C H F Y M E D L I D E Y L S F T K	731
Sc_Est2p	625	I R E D G L F Q G S S I L S A P I V D L V Y D D L L E F Y S E F K A	657
Ea_p123	730	K Q T K G I P Q G L C V S S I L S S F Y Y A T L E E S S L G F L R	762
Sp_Tip1p	732	K K G - - - - - S V L L R V V D D F L F I T V N K K D A K K	756
Sc_Est2p	658	S P S Q D - - - - - T L I L K L A D D F L I I S T D Q Q Q V I N	684
Ea_p123	763	D E S M N P E N P N V N L L M R L T D Y L L I T T Q E N N A V L	795
Sp_Tip1p	757	F L N I S L R G F E K H N F S T S L E K T V I N F E N S N G - - -	786
Sc_Est2p	685	I K K L A M G G F Q K Y N A K A N R D K I L A V S S Q S D - - -	713
Ea_p123	796	F I E K L I N V S R E N G F K F N M K K L Q T S F P L S P S K F A	828
Sp_Tip1p	787	- - - I I N N T F F N E S K K R M P F F G F S V N M R S L D T L L	816
Sc_Est2p	714	- - - D D T V I Q F C A - - M H I F V K E L E V W K H S S T M	739
Ea_p123	829	K Y G M D S V E E Q N I V Q D Y C D W I G I S I D M K T L A L M P	861
Sp_Tip1p	817	A C P K I D E A L F N S T S V E L T K H M G K S F F Y K I L R S S	849
Sc_Est2p	740	N N F H I R S K S S K G I F R S L I A L F N T R I S Y K T I D T N	772
Ea_p123	862	N I N L R I E G I L C T L N L N M Q T K K A S M W L K K L K S F	894

FIG. 42D



4

Sp_Tip1p	850	LASFAQVFI NSTNTV LMNN	DIT MQIDH ITHYFRKT	I HNSK VVKN T	SCCN SEC FANKT	I YR L G F	Y S MC MR	882
Sc_Est2p	773	LN S T	NT V L	DK H	SEC	-	-	793
Ea_p123	895	LN M N	NN I	DK H	DK T	LN K L	I S GG Y K	927

Sp_Tip1p	883	A Q A Y L K R M K D I F   P Q R M F I T D L L N V I G R K   I W K K	915
Sc_Est2p	794	- - Y K S A F K D L S I N - - V T Q N M Q F H S F L Q R   I E M	821
Ea_p123	928	Y M Q C A K E Y K D H E K K N L A M S S M I D L E V S K   - Y S V	960

Sp_Tip1p	916	LAE I L G Y T S R R F L S S A E V K W L F C L G M R D G L K P S	948
Sc_Est2p	822	T V S G C P I I K C D P L I E Y E V R F T I L N G F L	854
Ea p123	961	T R A F F K Y L V C N I K D T I F G E E H Y P D F F L S T L K H F	993

Sp_Tip1p	949	F K Y H P C F E Q L I Y Q F Q S L T D L I K P L R P V L R Q V L F	981
Sc_Est2p	855	T S - - - - - K F K D N I I L R K E I Q H L Q A Y I Y	877
Ea p123	994	I E I F S - - - T K K Y I F N R V C M I L K A K E A K L K S D Q C	1024

Sp_Tip1p	982	L H R R I A D -	988
Sc_Est2p	878	I Y I H I V N -	884
Ea_p123	1024	Q S L I Q Y D A	103

FIG. 42E



42/78

B.

Sp_Tip1p	1	- - - - -	M T E H H T P K S R I L R F L E N Q Y Y L C T	24
Sc_Est2p	1	- - - - -	- - - - - M K I L F E F	7
Ea_p123	1	M E V D V D N Q A D N H G I H S A L K T C E E I K E A K T L Y S W	33	
Sp_Tip1p	25	L N D Y Y Q L V L R G S P A S S Y S N I C E R L R S D V Q T S F S	57	
Sc_Est2p	8	Q D K L D I D L Q T N - - S T Y K - - E N L K C G H F N G L D	35	
Ea_p123	34	Q K V I R C R N Q S Q - - S H Y K - - D L E D I K I F A Q T N	61	
Sp_Tip1p	58	F L H S T V V G F D S K P D E G V Q F S S P K C S Q S E L [ I A N	90	
Sc_Est2p	36	E I L T C F A L P N S R - K I A L P C L P G D L S H K A V [ I D H	67	
Ea_p123	62	V A T P R D Y N E E D F K V I A R K E V F S T G L M I E L [ I D K	94	
Sp_Tip1p	91	Y Y K Q M F D E S F E R R R - N L L M K [ G F S M N H E D F R A M H	122	
Sc_Est2p	68	C I Y L L T G E L Y N - - N V L T F G Y K I A R N E D - - -	93	
Ea_p123	95	C L V E L L S S S D V S D R Q K L Q C F G F Q L K G N Q - - -	122	
Sp_Tip1p	123	V N G V Q N D L V S T F P N Y L I S I L E S K N [ W Q L L E I I G	155	
Sc_Est2p	94	- - - V N N S L F C H S A N V N V T L L K G A A W K M F H S L V G	123	
Ea_p123	123	- - - L A K T H L L T A L S T Q K Q Y F F Q D E W N Q V R A M I G	152	
Sp_Tip1p	156	S D A M H Y [ L S K G S I F E A L P N D N Y L Q I S G I P L F K N	188	
Sc_Est2p	124	T Y A F V D L L I N Y T V I Q F N - G Q F F T Q I V G N R C N E P	155	
Ea_p123	153	N E L F R H L Y T K Y L I F Q R T S E G T L V Q F C G N N V F D H	185	
Sp_Tip1p	189	N V F E E T V S K K R K R T I E T S I T Q N - - - K S A R K E V S	218	
Sc_Est2p	156	H L P P K W V Q - - R S S S S A T A A Q I - - - K Q L T E P V T	183	
Ea_p123	186	L K V N D K F D K - K Q K G G A A D M N E P R C C S T C K Y N V K	217	

FIG. 42F

+

+



43/78

B.

Sp_Tip1p	219	WNSISSISRFSSIFYRSSYKKFKQDLYFNLHSICD	251
Sc_Est2p	184	N-----KQFLHKLNINSSFFP	200
Ea_p123	218	NEK - DHFLNNINVPNNMKSRTRIFYCTHFN	248
Sp_Tip1p	252	RNTVHWWLQWIFPRQFGLINAFAQVKQLVIP	284
Sc_Est2p	201	-----YSKILPSSS-----SIKKLTDLREAIFP	223
Ea_p123	249	R-----NNQFFKKHEFVSNKNNISAMDRAQT	275
Sp_Tip1p	285	VS-----QSTVVPKRLLKVYPLIEQTAKRLHRI	313
Sc_Est2p	224	TN-----LVKIPQRLKVRINLTQKLLKRHKRLN	252
Ea_p123	276	FTNIFRFNRIRKKLKDVKIEKIAYMLEKVKDFN	308
Sp_Tip1p	314	LSKVVYNHYCPYID-THDDEKILSYSLKPNQ	342
Sc_Est2p	253	YVSILNSICPPLEGTVDDLSHLSRQSPKER	282
Ea_p123	309	FNYYLTKSCPPLPENWRERKQKIELINKTREEK	341
Sp_Tip1p	343	-----VFAFLRSILVRYFPKLI	359
Sc_Est2p	283	-----VLKFIIVILQKLLPQEM	299
Ea_p123	342	SKYYPEELFSYTTDNKCVTQFINEFFYNILPKDF	374
Sp_Tip1p	360	WGNQRIFEIILKDLETFLLKLSRYESFSLHYLMS	392
Sc_Est2p	300	FGSKKKNGKIKNLLSPLNGYLPFDSSLK	332
Ea_p123	375	LTG-RNRKNFQKKVKKYVELNKHELIHKNLLLE	406
Sp_Tip1p	393	NIKISEI <del>EWL</del> VLGKRSNAKMC <del>LSDFE</del> KRKQIFA	425
Sc_Est2p	333	KLRLKDFRWLFIS-----DIWFTKHNFENLNQLAI	362
Ea_p123	407	KINTREISWMQVETS-AKHFFYYFDHEN-IYVLW	437

FIG. 42G

+



44/78

B.

Sp_Tip1p	426	EF I YWL YNSF I I PIL QSF Y I TE SSDLRNRTVY	458
Sc_Est2p	363	CF I SWL F RQL I PK I I QTF Y CTE I SST VT - I VY	394
Ea_p123	438	KLLRW I FEDL V VSL I RCFFY VTEQQKSYSKTY	470
Sp_Tip1p	459	FRKD I WKLLCRPF ITSMKWEAFEKINENNVRMD	491
Sc_Est2p	395	FRHDTWNKLITPF I VEYFKTYLVENNVCRNHNS	427
Ea_p123	471	YRKNIWDVIMKMSIADLKKETLAEVQEKEVEEW	503
Sp_Tip1p	492	TQKTTLP PAV I RLLPKK - NTFRL I TNLRKRF L	522
Sc_Est2p	428	YTL SNFNHSKMR I PKK SNN EFR I A I PCRGAD	460
Ea_p123	504	KKSLGFA PGKLRLIPKK - TTFRP I MTFNKKIV	534
Sp_Tip1p	523	IKMGSNKKMLVSTNQTLRPVAS I LKHLLINE - -	552
Sc_Est2p	461	EEE - FT I YKENHKNAI QPTQK I LEYLRNKRPT	491
Ea_p123	535	NSD - - RKTTLTNTKLLNSHLMMLKTLKNR - MF	564
Sp_Tip1p	553	ESSGIPFNL EYMKLLTFKKQRL LKKH R M F G R - KK	584
Sc_Est2p	492	SFTK I YSPPTQ I ADRIKEFKQRL LKKFNNVLP E L	524
Ea_p123	565	KDPFGFA VFN YDDVMKKYEEFVCKWKQVGQPKL	597
Sp_Tip1p	585	YF VRI D I KSCYDR I KQDLMFR I VKKKLKDPE - F	616
Sc_Est2p	525	YF MKFD V KSCYDS I PRMECMR I LKDALKNENG F	557
Ea_p123	598	FFATMD I EKCYDS V NREKLSTFLKTTKLSSDF	630
Sp_Tip1p	617	VIRKYAT I HATSDRATKN - - - - -	634
Sc_Est2p	558	FVR S Q Y F FNTNTG - - - - -	570
Ea_p123	631	WIMTAQ I LKRKN I VIDSKNFRKKEMKD YFRQK	663

FIG. 42H



45/78

B.

Sp_Tip1p	635	F V S E A F S Y F D M V P F E K V V Q L L S - - M K T S D T [ ] F V	665
Sc_Est2p	571	- - - - - V L K L F N V V N A S R - - V P K P Y E L Y -	591
Ea_p123	664	F Q K I A L E G G Q Y P T L F S V L E N E Q N D L N A K K T L V	696
Sp_Tip1p	666	D F V D Y W T K S S S E I F K M L K E H L S G H I V K I G N S Q Y	698
Sc_Est2p	592	D N V R T V H L S N Q D V I N V V E M E I F K T A L W V E D K C Y	624
Ea_p123	697	E A K Q R N Y F K K D N L L Q P V I N I C Q Y N Y I N F N G K F Y	729
Sp_Tip1p	699	L Q K V [ ] G I P [ ] Q G S I L S F L C H F Y M E D L I D E Y L S F T K	731
Sc_Est2p	625	I R E D G L F Q G S S L S A P I V D L V Y D D L L E F Y S E F K A	657
Ea_p123	730	K Q T K G I P Q G L C V S S I L S S F Y Y A T L E E S S L G E L R	762
Sp_Tip1p	732	K K G - - - - S V L L R V V [ ] D D F [ ] F I T V N K K D A K K	756
Sc_Est2p	658	S P S Q D - - - - T L I L K L A D D F [ ] L I I S T D Q Q Q V I N	684
Ea_p123	763	D E S M N P E N P N V N L L M R L T D D Y L L I T T Q E N N A V L	795
Sp_Tip1p	757	F L N L S L R G F E K H N F S T S L E K T V I N F E N S N G - - -	786
Sc_Est2p	685	I K K L A M G G F Q K Y N A K A N R D K I L A V S S Q S D - - -	713
Ea_p123	796	F I E K L I N V S R E N G F K F N M K K L Q T S F P L S P S K F A	828
Sp_Tip1p	787	- - - I I N N T F F N E S K K R M P F F G F S V N M R S L D T L L	816
Sc_Est2p	714	- - - D D T V I Q F C A - - M H I F V K E L E V V W K H S S T M	739
Ea_p123	829	K Y G M D S V E E Q N I V Q D Y C D W I G I S I D M K T L A L M P	861
Sp_Tip1p	817	A C P K I D E A L F N S T S V E L T K H M G K S F F Y K I L R S S	849
Sc_Est2p	740	N N F H I R S K S S K G I F R S L I A L F N T R I S Y K T I D T N	772
Ea_p123	862	N I N L R I E G I L C T L N L N M Q T K K A S M W L K K K L K S F	894

FIG. 42/

+



46/78

B.

Sp_Tip1p	850	[LASFAQVFIGIDITHNSKFNSSCCNIVRLGYSMCMR	882
Sc_Est2p	773	[LNSTNTVLMQIDHVVKNISEC	793
Ea_p123	895	[LMNNITHYFRKTITTEDFANKTLNKLFI SGGYK	927
Sp_Tip1p	883	AQAYLKR[MKD]FIPQRMFI TDLLNVIGRK[WK	915
Sc_Est2p	794	- - - YKSAFKDL SIN - VTQNMQFHSFLQR[EM	821
Ea_p123	928	YMQCAKEY[KDHFKKNLAMSSMIDLEVSKI[YSV	960
Sp_Tip1p	916	LAEILGYTSRRFLSSAEVKWLFCLGMRDGLKPS	948
Sc_Est2p	822	TVSGCPITKCDPLIEYEVRFTILNGFLESLSSN	854
Ea_p123	961	TRAFFFKYLVVCNIKDTIFGEEHYPDFFLSTLKH	993
Sp_Tip1p	949	FKYHPCFEQLIYQFQSLTDLIKPLRPVLRQVLF	981
Sc_Est2p	855	TS - - - KFKDNILLRKEIQHLQAYIY	877
Ea_p123	994	IEIFS - - - TKKYIFNRVCMIKAKEAKLKS DQC	1023
Sp_Tip1p	982	LHRRIAD -	988
Sc_Est2p	878	IYIHIVN -	884
Ea_p123	1024	QSLIQYDA	1031

FIG. 42J

+



47/78

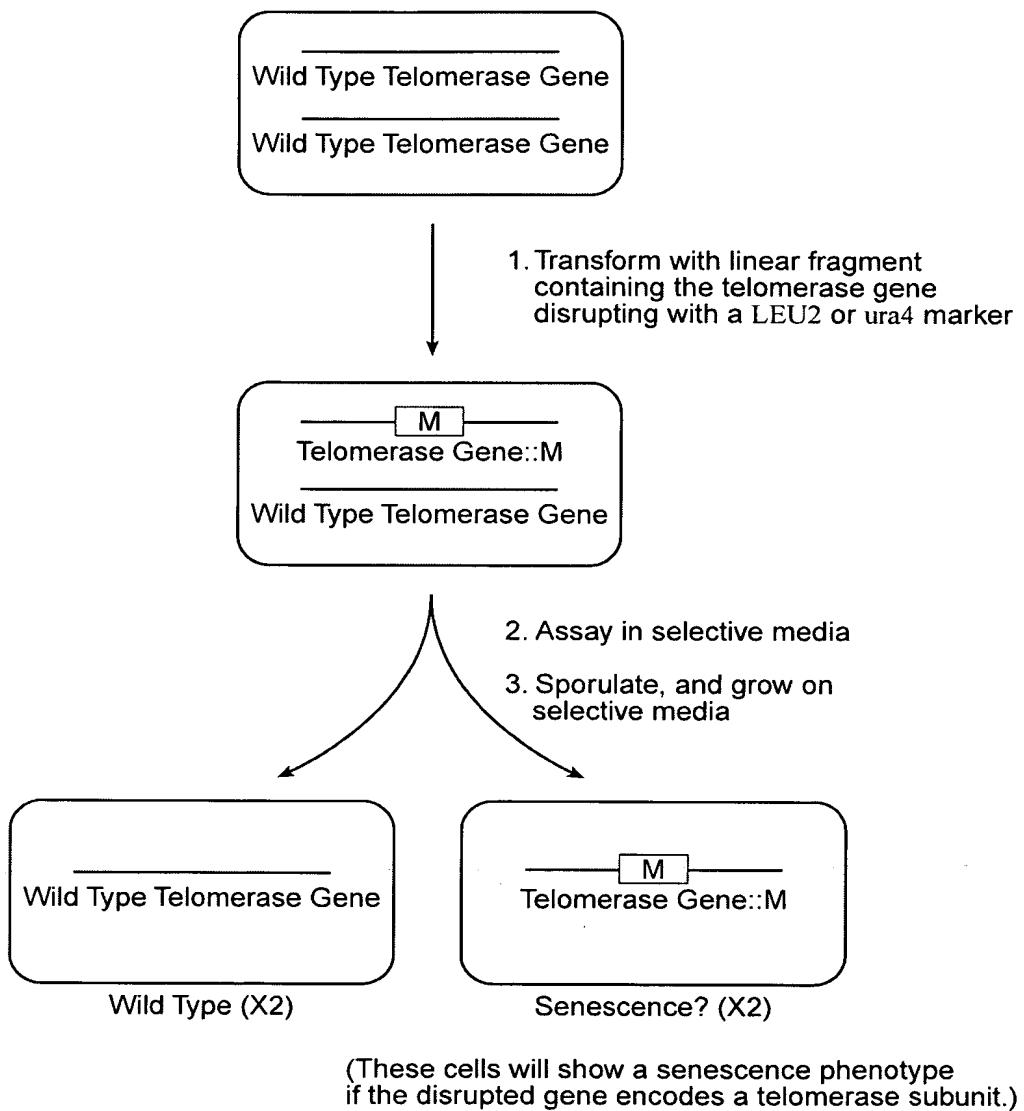


FIG. 43

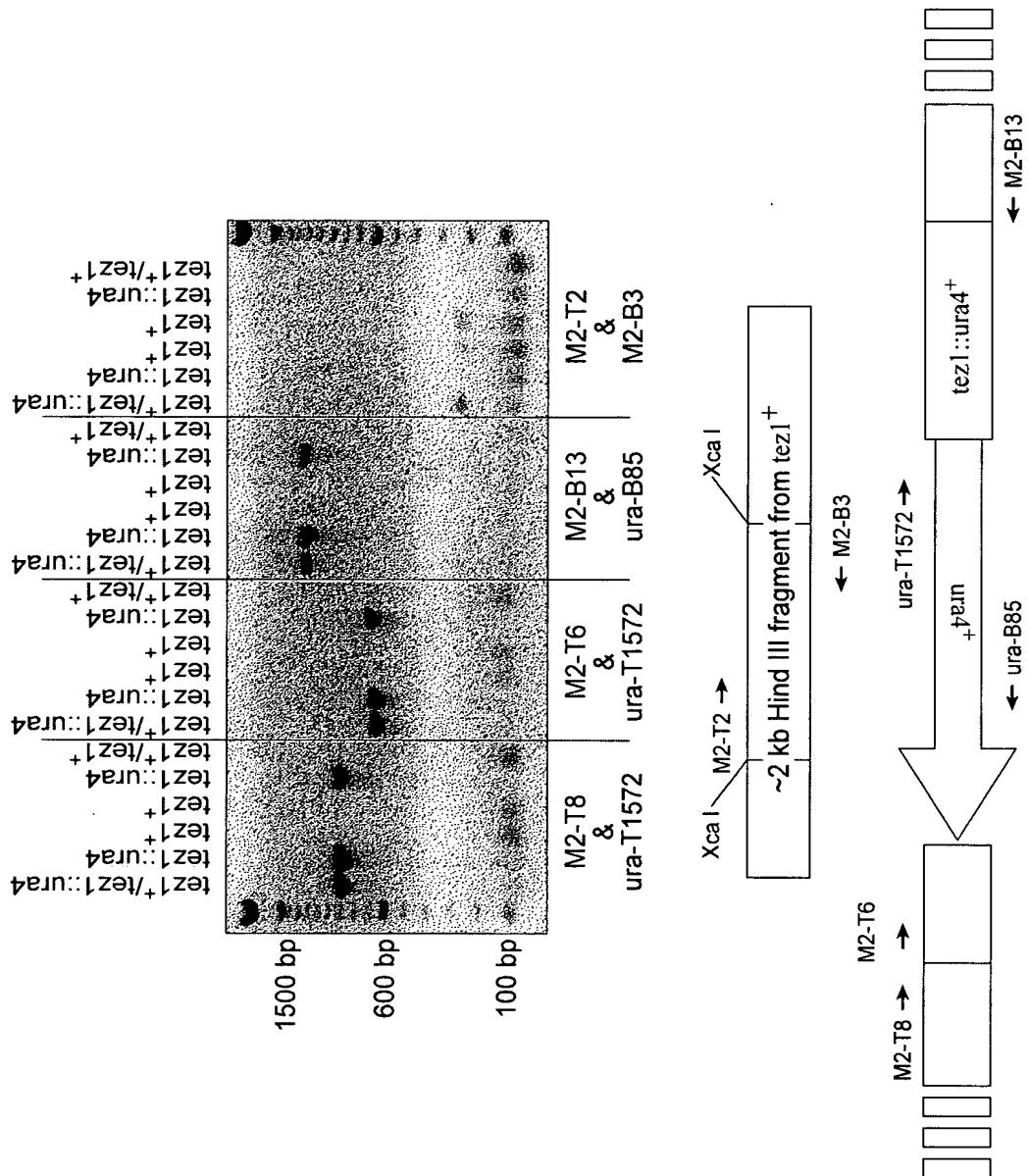


FIG. 44

+



49/78

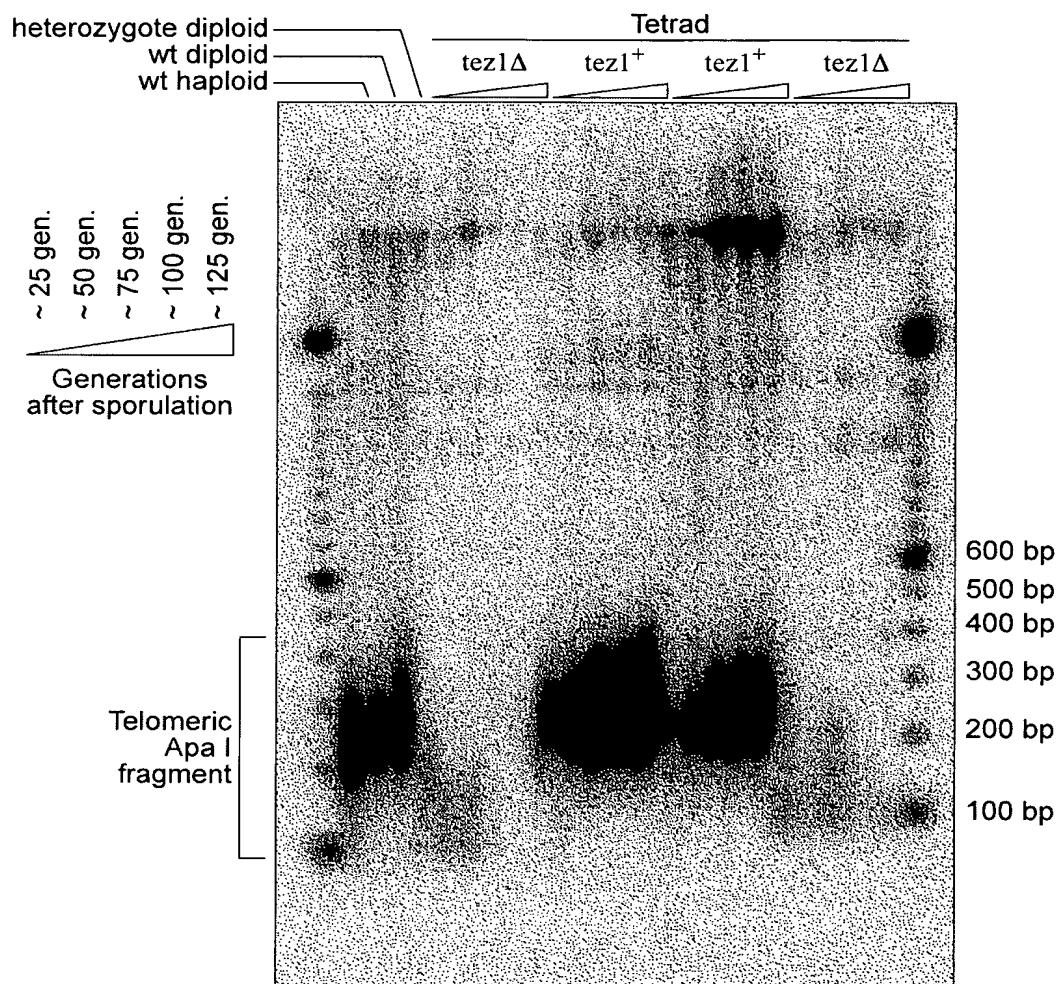


FIG. 45

+



50/78

FIG. 46A



51/78

1470 GAT CTC GTT TCT ACT TTT CCT AAT TAC CTT ATA TCT CTT GAG TCA AAA AAT TGG CAA 1529  
 129 D L V S T F P N Y L I S I L E S K N W Q 148  
 1530 CTT TTG TTA GAA AT gtaaataccggtaaaggatgtgcacttacaagaactgacaatgtatag T ATC GGC 1601  
 149 L L E I I G 155  
 1602 AGT GAT GCC ATG CAT TAC TTA TCC AAA GGA AGT ATT TTT GAG GCT CTT CCA AAT GAC 1661  
 156 S D A M H Y L L S K G S I F E A L P N D 175  
 1662 AAT TAC CTT CAG ATT TCT GGC ATA CCA CTT TTT AAA AAT AAT GTG TTT GAG GAA ACT GTG 1721  
 176 N Y L Q I S G I P L F K N N V F E E T V 195  
 1722 TCA AAA AGA AAG CGA ACC ATT GAA ACA TCC ATT ACT CAA AAT AAA AGC GCC CGC AAA 1781  
 196 S K K R K R T I E T S I T Q N K S A R K 215  
 1782 GAA GTT TCC TGG AAT AGC ATT TCA ATT AGT AGG TTT AGC ATT TTT TAC AGG TCA TCC TAT 1841  
 216 E V S W N S I S R F S I F Y R S S Y 235  
 1842 AAG AAG TTT AAG CAA G gtaactaatactgttatcccttataactaatttag AT CTA TAT TTT AAC 1907  
 236 K F K Q D L Y F N 245  
 1908 TTA CAC TCT ATT TGT GAT CGG AAC ACA GTA CAC ATG TGG CTT CAA TGG ATT TTT CCA AGG 1967  
 246 L H S I C D R N T V H M W L Q W I F P R 265  
 1968 CAA TTT GGA CTT ATA AAC GCA TTT CAA GTG AAG CAA TTG CAC AAA GTG ATT CCA CTG GTA 2027  
 266 Q F G L I N A F Q V K Q L H K V I P L V 285  
 2028 TCA CAG AGT ACA GTT GTG CCC AAA CGT CTC CTA AAG GTA TAC CCT TTA ATT GAA CAA ACA 2087  
 286 S Q S T V V P K R L L K V Y P L I E Q T 305  
 2088 GCA AAG CGA CTC CAT CGT ATT TCT CTA TCA AAAA GTT TAC AAC CAT TAT TGC CCA TAT ATT 2147  
 306 A K R L H R I S L S K V Y N H Y C P Y I 325  
 2148 GAC ACC CAC GAT GAT GAA AAA ATC CTT AGT ATT TCC TTA AAG CCG AAC CAG GTG TTT GCG 2207  
 326 D T H D E K I L S Y S L K P N Q V F A 345  
 2208 TTT CGA TCC ATT CTT GTT CGA GTG TTT CCT AAA TTA ATC TGG GGT AAC CAA AGG ATA 2267  
 346 F L R S I L V R V F P K L I W G N Q R I 365

F/G. 46B



52/78

2268 TTT GAG ATA ATA TTA AAA G gtattgtataaaatttaccactaacgattttaccag AC CTC GAA ACT 2336  
 366 F E I I L K D L E T 375  
  
 2337 TTC TTG AAA TTA TCG AGA TAC GAG TCT TTT AGT TTA CAT TAT TTA ATG AGT AAC ATA AAG 2396  
 376 F L K L S R Y E S F S L H Y L M S N I K 395  
  
 2397 gtaatatgccaaattttaccattaaacaatcg ATT TCA GAA ATT GAA TGG CTA GTC CTT GGA 2465  
 396  
  
 2466 AAA AGG TCA AAT GCG AAA ATG TGC TTA AGT GAT TTT GAG AAA CGC AAG CAA ATA TTT GCG 2525  
 406 K R S N A K M C L S D F E K R K Q I F A 425  
  
 2526 GAA TTC ATC TAC TGG CTA TAC AAT TCG TTT ATA ATA CCT ATT TTA CAA TCT TTT TTT TAT 2585  
 426 E F I Y W L Y N S F I I P I L Q S F F Y 445  
  
 2586 ATC ACT GAA TCA AGT GAT TTA CGA AAT CGA ACT GTT TAT TTT AGA AAA GAT ATT TGG AAA 2645  
 446 I T E S S D L R N R T V Y F R K D I W K 465  
  
 2646 CTC TTG TGC CGA CCC TTT ATT ACA TCA ATG AAA ATG GAA GCG TTT GAA AAA ATA AAC GAG 2705  
 466 L C R P F I T S M K M E A F E K I N E 485  
  
 2706 gtattttaaagtattttgcaaaaaggctaataattttcg AAC AAT GTT AGG ATG GAT ACT CAG AAA ACT 2775  
 486 N N V R M D T Q K T 495  
  
 2776 ACT TTG CCT CCA GCA GGT ATT CGT CTA TTA CCT AAG AAG AAT ACC TTT CGT CTC ATT ACG 2835  
 496 T L P P A V I R L L P K K N T F R L I T 515  
  
 2836 AAT TTA AGA AAA AGA TTC TTA ATA AAG gtattaaattttttggcatcaatgtactttcaattctattata 2906  
 516 N L R K R F L I K  
  
 2907 tttagcag ATG GGT TCA AAC AAA AAA ATG TTA GTC AGT ACG AAC CAA ACT TTA CGA CCT GTG 2967  
 525 M G S N K K M L V S T N Q T L R P V 542  
  
 2968 GCA TCG ATA CTG AAA CAT TTA ATC AAT GAA GAA AGT AGT GGT ATT CCA TTT AAC TTG GAG 3027  
 543 A S I L K H L I N E E S S G I P F N L E 562  
  
 3028 GTT TAC ATG AAG CTT CTT ACT TTT AAG AAG GAT CTT AAG CAC CGA ATG TTT GG gtaat 3088  
 563 V M K L L T F K D L K H R M F G 581

FIG. 46C

+



53/78

FIG. 46D

+



54/78

FIG. 46E

1



FIG. 46F



56/78

1  
 met ser val tyr val val glu leu leu  
 GCCAAGTTCCCTGCACTGGCTG      ATG AGT GTG TAC GTC GTC GAG CTG CTC  
  
 10  
 arg ser phe phe tyr val thr glu thr thr phe gln lys asn arg  
 AGG TCT TTC TTT TAT GTC ACG GAG ACC ACG TTT CAA AAG AAC AGG  
  
 20  
 leu phe phe tyr arg lys ser val trp ser lys leu gln ser ile  
 CTC TTT TTC TAC CGG AAG AGT GTC TGG AGC AAG TTG CAA AGC ATT  
  
 30  
 gly ile arg gln his leu lys arg val gln leu arg glu leu ser  
 GGA ATC AGA CAG CAC TTG AAG AGG GTG CAG CTG CGG GAG CTG TCG  
  
 40  
 glu ala glu val arg gln his arg glu ala arg pro ala leu leu  
 GAA GCA GAG GTC AGG CAG CAT CGG GAA GCC AGG CCC GCC CTG CTG  
  
 50  
 thr ser arg leu arg phe ile pro lys pro asp gly leu arg pro  
 ACG TCC AGA CTC CGC TTC ATC CCC AAG CCT GAC GGG CTG CGG CCG  
  
 60  
 ile val asn met asp tyr val val gly ala arg thr phe arg arg  
 ATT GTG AAC ATG GAC TAC GTC GTG GGA GCC AGA ACG TTC CGC AGA  
  
 70  
 glu lys ala glu arg leu thr ser arg val lys ala leu phe  
 GAA AAG ARG GCC GAG CGT CTC ACC TCG AGG GTG AAG GCA CTG TTC  
  
 80  
 ser val leu asn tyr glu arg ala arg arg pro gly leu leu gly  
 AGC GTG CTC AAC TAC GAG CGG GCG CGG CGC CCC GGC CTC CTG GGC  
  
 90  
 100  
 ala ser val leu gly leu asp asp ile his arg ala trp arg thr  
 GCC TCT GTG CTG GGC CTG GAC GAT ATC CAC AGG GCC TGG CGC ACC  
  
 110  
 phe val leu arg val arg ala gln asp pro pro pro glu leu tyr  
 TTC GTG CTG CGT GTG CGG GCC CAG GAC CCG CCT GAG CTG TAC  
  
 120  
 phe val lys val asp val thr gly ala tyr asp thr ile pro gln  
 TTT GTC AAG GTG GAT GTG ACG GGC GCG TAC GAC ACC ATC CCC CAG  
  
 130  
 140  
 asp arg leu thr glu val ile ala ser ile ile lys pro gln asn  
 GAC AGG CTC ACG GAG GTC ATC GCC AGC ATC AAA CCC CAG AAC  
  
 150  
 160  
 thr tyr cys val arg arg tyr ala val val gln lys ala ala met  
 ACG TAC TGC GTG CGT CGG TAT GCC GTG GTC CAG AAG GCC GCC ATG

FIG. 47A



57/78

210  
 gly thr ser ala arg pro ser arg ala thr ser tyr val gln cys  
 GGC ACG TCC GCA AGG CCT TCA AGA GCC ACG TCC TAC GTC CAG TGC

220 230  
 gln gly ile pro gln gly ser ile leu ser thr leu leu cys ser  
 CAG GGG ATC CCG CAG GGC TCC ATC CTC TCC ACG CTG CTC TGC AGC

240  
 leu cys tyr gly asp met glu asn lys leu phe ala gly ile arg  
 CTG TGC TAC GGC GAC ATG GAG AAC AAG CTG TTT GCG GGG ATT CGG

250 260  
 arg asp gly leu leu leu arg leu val asp asp phe leu leu val  
 CGG GAC GGG CTG CTC CTG CGT TTG GTG GAT GAT TTC TTG TTG GTG

270  
 thr pro his leu thr his ala lys thr phe leu arg thr leu val  
 ACA CCT CAC CTC ACC CAC GCG AAA ACC TTC CTC AGG ACC CTG GTC

280 290  
 arg gly val pro glu tyr gly cys val val asn leu arg lys thr  
 CGA GGT GTC CCT GAG TAT GGC TGC GTG GTG AAC TTG CGG AAG ACA

300  
 val val asn phe pro val glu asp glu ala leu gly gly thr ala  
 GTG GTG AAC TTC CCT GTA GAA GAC GAG GCC CTG GGT GGC ACG GCT

310 320  
 phe val gln met pro ala his gly leu phe pro trp cys gly leu  
 TTT GTT CAG ATG CCG GCC CAC GGC CTA TTC CCC TGG TGC GGC CTG

330  
 leu leu asp thr arg thr leu glu val gln ser asp tyr ser ser  
 CTG CTG GAT ACC CGG ACC CTG GAG GTG CAG AGC GAC TAC TCC AGC

340 350  
 tyr ala arg thr ser ile arg ala ser leu thr phe asn arg gly  
 TAT GCC CGG ACC TCC ATC AGA GCC AGT CTC ACC TTC AAC CGC GGC

360  
 phe lys ala gly arg asn met arg arg lys leu phe gly val leu  
 TTC AAG GCT GGG AGG AAC ATG CGT CGC AAA CTC TTT GGG GTC TTG

370 380  
 arg leu lys cys his ser leu phe leu asp leu gln val asn ser  
 CGG CTG AAG TGT CAC AGC CTG TTT CTG GAT TTG CAG GTG AAC AGC

390  
 leu gln thr val cys thr asn ile tyr lys ile leu leu leu gln  
 CTC CAG ACG GTG TGC ACC AAC ATC TAC AAG ATC CTC CTG CTG CAG

400 410  
 ala tyr arg phe his ala cys val leu gln leu pro phe his gln  
 GCG TAC AGG TTT CAC GCA TGT GTG CTG CAG CTC CCA TTT CAT CAG

FIG. 47B



58/78

420

gln val trp lys asn pro his phe ser cys ala ser ser leu thr  
 CAA GTT TGG AAG AAC CCA CAT TTT TCC TGC GCG TCA TCT CTG ACA

430

440

arg leu pro leu leu leu his pro glu ser gln glu arg arg asp  
 CGG CTC CCT CTG CTA CTC CAT CCT GAA AGC CAA GAA CGC AGG GAT

450

val ala gly gly gln gly arg arg arg pro ser ala leu arg gly  
 GTC GCT GGG GGC CAA GGG CGC CGC CGG CCC TCT GCC CTC CGA GGC

460

470

arg ala val ala val pro pro ser ile pro ala gln ala asp ser  
 CGT GCA GTG GCT GTG CCA CCA AGC ATT CCT GCT CAA GCT GAC TCG

480

thr pro cys his leu arg ala thr pro gly val thr gln asp ser  
 ACA CCG TGT CAC CTA CGT GCC ACT CCT GGG GTC ACT CAG GAC AGC

490

500

pro asp ala ala glu ser glu ala pro gly asp asp ala asp cys  
 CCA GAC GCA GCT GAG TCG GAA GCT CCC GGG GAC GAC GCT GAC TGC

510

pro gly gly arg ser gln pro gly thr ala leu arg leu gln asp  
 CCT GGA GGC CGC AGC CAA CCC GGC ACT GCC CTC AGA CTT CAA GAC

520

530

his pro gly leu met ala thr arg pro gln pro gly arg glu gln  
 CAT CCT GGA CTG ATG GCC ACC CGC CCA CAG CCA GGC CGA GAG CAG

540

thr pro ala ala leu ser arg arg ala tyr thr ser gln gly gly  
 ACA CCA GCA GCC CTG TCA CGC CGG GCT TAT ACG TCC CAG GGA GGG

550

560

arg gly gly pro his pro gly leu his arg trp glu ser glu ala  
 AGG GGC GGC CCA CAC CCA GGC CTG CAC CGC TGG GAG TCT GAG GCC

564

OP  
 TGA GTGAGTGTGTTGGCCGAGGCCTGCATGTCCGGCTGAAGGCTGAGTGTCCGGCTGAGGC  
 CTGAGCGAGTGTCCAGCCAAGGGCTGAGTGTCCAGCACACCTGCGTTTCACTTCCCCAC  
 AGGCTGGCGTTGGTCCACCCCAAGGGCCAGCTTTCCCTACCAGGAGCCGGCTTCCACT  
 CCCCCACATAGGAATAGTCCATCCCCAGATTGCCATTGTTCACCCCTGCCCTGCCTTCC  
 TTTGCCTTCCACCCCCACCATTAGGTGGAGACCTGAGAAGGACCCCTGGAGCTTGGG  
 AATTGGAGTGACCAAGGTGTGCCCTGTACACAGGGAGGACCCCTGCACCTGGATGGGG  
 GTCCCTGTGGGTCAAATTGGGGAGGTGCTGTGGAGTAAAATACTGAATATATGAGTT  
 TTTCAGTTTGGAAAAAAAAAAAAAAAAAAAAAA

FIG. 47C



59/78

Motif -1

Ep p123	...LVVSLIRCFYYVTEQQKSYSKT...
Sp Tez1	...FTIPILOQSFYITESSDLRNRT...
Sc Est2	...LIPKIIQTFYYCTEISSTVTIV...
Hs TCP1	...YVVELLRSFFYVTETTFQKNRL...
consensus	FFY TE

Motif 0

Ep p123	p hhh K hR h	K
Sp Tez1	...KSLGFAPGKLRLIPKKT--TFRPIMTFNKKIV...	R
Sc Est2	...QKTTLPPAVIRLLPKKN--TFRLITNLRKRLF...	
Hs TCP1	...TLSNFNHSKMRIIPKKSNNFRIIAIPCRGAD...	
consensus	...ARPALLTSRLRFIPKPD--GLRPIVNMDYVVG...	
	R PK R I	

Motif A

Ep p123	h hDh GY h	AF
Sp Tez1	...PKLFFATMDIEKCYDSVNREKLSTFLK...	
Sc Est2	...RKKYFVRIDIKSCYDRIKQDLMFRIVK...	
Hs TCP1	...PELYFMKFDVKSCYDSIPRMECMRILK...	
consensus	...PELYFVKVDVTGAYDTIPQDRLTEVIA...//...	
	F D YD	

Motif B

Ep p123	hPQG pS hh	
Sp Tez1	...NGKFYKQTKGIPQGLCVSSILSSFYA...	
Sc Est2	...GNSQYLQKVGIPQGSILSSFLCHFYME...	
Hs TCP1	...EDKCYIREDGLFQGSSLSAPIVDLVYD...	
consensus	...RATSYVQCQGIPQGSILSTLLCSLCYG...	
	G QG S	

Motif C

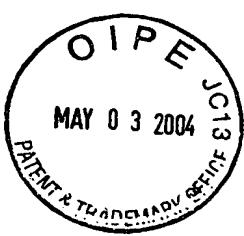
Ep p123	h F DD hhh	Y
Sp Tez1	...PNVNLLMRLTDDYLLITTQENN...	
Sc Est2	...KKGSVLLRVVDDFLFITVNKKD...	
Hs TCP1	...SQDTLILKLADDFLIISTDQQQ...	
consensus	...RRDGLLLRLVDDFLLVTPHLTH...	
	DD L	

Motif D

Ep p123	Gh h CK	
Sp Tez1	...NVSRENGFKFNMKKL...	
Sc Est2	...LNLSLRGFEKHNFST...	
Hs TCP1	...KKLAMGGFQKYNAKA...	
consensus	...LRTLVRGVPEYGCVV...	
	G	

FIG. 48

+



60/78

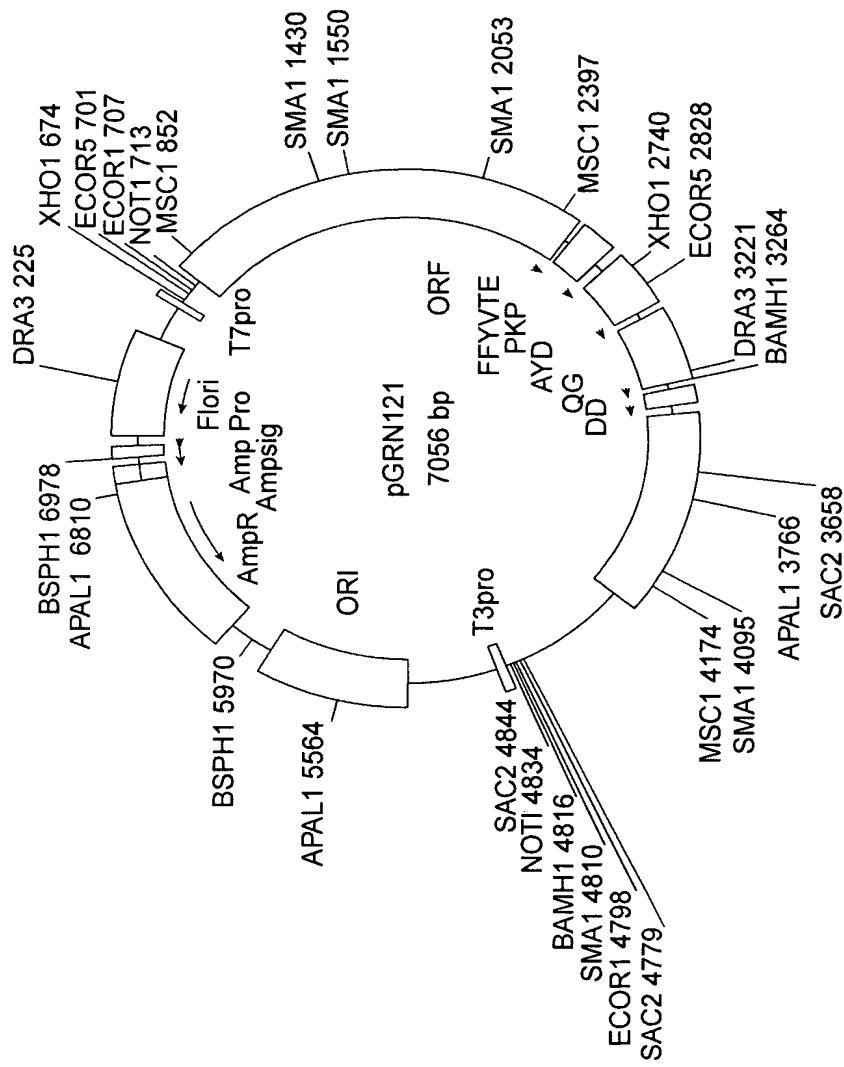


FIG. 4.9

+



61/78

1 GCAGCGCTGC GTCCTGCTGC GCACGTGGGA AGCCCTGGCC CCGGCCACCC  
 51 CGCGATGCC GCGCGCTCCC CGCTGCGGAG CGTGCGCTC CCTGCTGCGC  
 101 AGCCACTACC GCGAGGTGCT GCCGCTGGCC ACGTTCGTGC GGCGCCTGGG  
 151 GCCCCAGGGC TGGCGGCTGG TGCAGCGCGG GGACCCGGCG GCTTCCGCG  
 201 CGNTGGTGGC CCANTGCNTG GTGTGCGTGC CCTGGGANGN ANGGCNGCCC  
 251 CCCGCCGCC CCTCCCTCCG CCAGGTGTCC TGCTGAANG ANCTGGTGGC  
 301 CCGAGTGCTG CANANGCTGT GCGANCGCGG CGCGAANAAAC GTGCTGGCCT  
 351 TCGGCTTCGC GCTGCTGGAC GGGGCCCGCG GGGGCCCCCG CGAGGCCCTTC  
 401 ACCACCAGCG TGCGCAGCTA CCTGCCAAC ACGGTACCG ACGCAC TGCG  
 451 GGGGAGCGGG GCGTGGGGC TGCTGCTGCG CGCGTGGGC GACGACGTGC  
 501 TGGTTCACCT GCTGGCACGC TGCGCNGNT TTGTGCTGGT GGNTCCCAGC  
 551 TCGGCCCTACC ANGTGCGG GCCGCCGCTG TACAGCTCG GCGCTGCNAC  
 601 TCAGGCCCGG CCCCCGCCAC ACGCTANTGG ACCCGAANGC GTCTGGGATC  
 651 CAACGGGCCT GGAACCATAG CGTCAGGGAG GCGGGGGTCC CCCTGGGCTG  
 701 CCAGCCCCGG GTGCGAGGAG GCGCGGGGGC AGTGCCAGCC GAAGTCTGCC  
 751 GTTGCCTAACAG AGGCCAGGGC GTGGCGCTGC CCCTGAGGCC GAGCGGACGC  
 801 CCGTTGGCA GGGGTCTGG GCCCACCCGG GCAGGACGCC TGGACCGAGT  
 851 GACCGTGGTT TCTGTGTTG TGTACCTGCC AGACCCGCCG AAGAACCCAC  
 901 CTCTTGGAG GGTGCGCTCT CTGGCACGCC CCACTCCCAC CCATCCGTGG  
 951 GCCGCCAGCA CCACGCCGGC CCCCCATCCA CATCGCGGCC ACCACGTCC  
 1001 GGGACACGCC TTGTCCCCCG GTGTACGCCG AGACCAAGCA CTTCCCTTAC  
 1051 TCCTCAGGCG ACAAGNACAC TGCGNCCCTC CTTCTACTC AATATATCTG  
 1101 AGGCCAGGCC TGACTGGCGT TCGGGAGGTT CGTGGAGACA NTCTTCTGG  
 1151 TTCCAGGCCCT TGGATGCCAG GATTCCCCGC AGGTTGCCCG GCCTGCCCA  
 1201 GCGNTACTGG CAAATGCCGC CCCTGTTCT GGAGCTGCTT GGGAAACACG  
 1251 CGCAGTGCCTC CTACGGGGTG TTCCTCAAGA CGCACTGCC GCTGCGAGCT  
 1301 GCGGTCACTC CAGCAGGCCG TGTCTGTGCC CGGGAGAACG CCCAGGGCTC  
 1351 TGTGGCGGCC CCCGAGGAGG AGGAACACAG ACCCCCGTCG CTTGGTGCAG  
 1401 CTGCTCCGCC AGCACAGCAG CCCCTGGCAG GTGTACGGCT TCGTGGGGC  
 1451 CTGCCTGCGC CGGCTGGTGC CCCCAGGCC CTGGGGCTCC AGGCACAACG  
 1501 AACGCCGCTT CCTCAGGAAC ACCAAGAAGT TCATCTCCCT GGGGAAGCAT  
 1551 GCCAAGCTCT CGCTGCAGGA GCTGACGTGG AAGATGAGCG TGCGGGACTG  
 1601 CGCTTGGCTG CGCAGGAGCC CAGGGTTGG CTGTGTTCCG GCCGCAGAGC  
 1651 ACCGTCTGCG TGAGGAGATC CTGGCAAGT TCCTGCACTG GCTGATGAGT  
 1701 GTGTACGTGCG TCGAGCTGCT CAGGTCTTTC TTTTATGTCA CGGAGACCAC  
 1751 GTTCAAAAG AACAGGCTCT TTTTCTACCG GAAGAGTGTG TGGAGCAAGT  
 1801 TGCAAAGCAT TGGAATCAGA CAGCACTTGA AGAGGGTGCA GCTGGGGAG  
 1851 CTGTCGAAG CAGAGGTCAG GCAGCATCGG GAAGCCAGGC CGCCCTGCT  
 1901 GACGTCCAGA CTCCGCTTCA TCCCCAAGCC TGACGGGCTG CGGCCGATTG  
 1951 TGAACATGGA CTACGTGCG GGAGCCAGAA CGTCCCGCAG AGAAAAGAGG  
 2001 GCCGAGCGTC TCACCTCGAG GGTGAAGGCA CTGTTCAGCG TGCTCAACTA  
 2051 CGAGCGGGCG CGGCGCCCCCG GCCTCTGG CGCTCTGTG CTGGGCTGG  
 2101 ACGATATCCA CAGGGCTGG CGCACCTTCG TGCTGCGTGT GCGGGCCAG  
 2151 GACCCGCCGC CTGAGCTGTA CTTTGTCAAG GTGGATGTGA CGGGCGCGTA  
 2201 CGACACCATC CCCCAGGACA GGCTCACCGGA GGTCACTGCC AGCATCATCA  
 2251 AACCCAGAA CACGTACTGC GTGCGTGGT ATGCCGTGGT CCAGAAGGCC  
 2301 GCCCATGGC ACGTCCGCAA GGCCTTCAAG AGCCACGTCT CTACCTTGAC  
 2351 AGACCTCCAG CCGTACATGC GACAGTTCGT GGCTCACCTG CAGGANAACA  
 2401 GCCCGCTGAG GGATGCCGTC GTCATCGAGC AGAGCTCCTC CCTGAATGAG  
 2451 GCCAGCAGTG GCCTCTTCGA CGTCTTCCCA CGCTTCATGT GCCACACGC

FIG. 50A

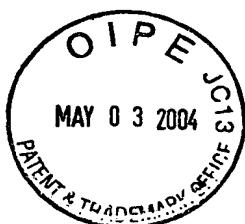
+



62/78

2501 CGTGCACATC AGGGCAAGT CCTACGTCCA GTGCCAGGGG ATCCCGCAGG  
2551 GCTCCATCCT CTCCACGCTG CTCTGCAGCC TGTGCTACGG CGACATGGAG  
2601 AACAAAGCTGT TTGCGGGGAT TCGGCAGGAC GGGCTGCTCC TGCCTTGGT  
2651 GGATGATTTC TTGTTGGTGA CACCTCACCT CACCCACGCG AAAACCTTCC  
2701 TCAGGACCCCT GGTCCGAGGT GTCCCTGAGT ATGGCTGCGT GGTGAACCTG  
2751 CGGAAGACAG TGGTGAACCTT CCCTGTAGAA GACGAGGCC TGGGTGGCAC  
2801 GGCTTTGTT CAGATGCCGG CCCACGGCCT ATTCCCCTGG TGCGGCCTGC  
2851 TGCTGGATAC CCGGACCCCTG GAGGTGCAGA GCGACTACTC CAGCTATGCC  
2901 CGGACCTCCA TCAGAGCCAG TCTCACCTTC AACCGCGGCT TCAAGGCTGG  
2951 GAGGAACATG CGTCGAAAC TCTTTGGGTT CTTGCGGCTG AAGTGTACA  
3001 GCCTGTTCT GGATTGCGAG GTGAACAGCC TCCAGACGGT GTGCACCAAC  
3051 ATCTACAAGA TCCTCTGCT GCAGGGTAC AGGTTTCACG CATGTGTGCT  
3101 GCAGCTCCCA TTTCATCAGC AAGTTTGGAA GAACCCCACA TTTTCTGC  
3151 GCGTCATCTC TGACACGGCC TCCCTCTGCT ACTCCATCCT GAAAGCCAAG  
3201 AACGCAGGGGA TGTCGCTGGG GGCCAAGGGC GCCGCCGGCC CTCTGCCCTC  
3251 CGAGGCCGTG CAGTGGCTGT GCCACCAAGC ATTCTGCTC AAGCTGACTC  
3301 GACACCGTGT CACCTACGTG CCACTCCTGG GGTCACTCAG GACAGCCCAG  
3351 ACGCAGCTGA GTCGGAAGCT CCCGGGGACG ACGCTGACTG CCCTGGAGGC  
3401 CGCAGCCAAC CCGGCACTGC CCTCAGACTT CAAGACCATC CTGGACTGAT  
3451 GGCCACCCGC CCACAGCCAG GCCGAGAGCA GACACCAGCA GCCCTGTCAC  
3501 GCCGGGCTCT ACGTCCCAGG GAGGGAGGGG CGGCCACAC CCAGGCCCGC  
3551 ACCGCTGGGA GTCTGAGGCC TGAGTGAGTG TTTGGCCGAG GCCTGCATGT  
3601 CCGGCTGAAG GCTGAGTGTG CGGCTGAGGC CTGAGCGAGT GTCCAGCAA  
3651 GGGCTGAGTG TCCAGCACAC CTGCCGTCTT CACTTCCCCA CAGGCTGGCG  
3701 CTCGGCTCCA CCCCAGGGCC AGCTTTCTT CACCAGGAGC CCGGCTTCCA  
3751 CTCCCCACAT AGGAATAGTC CATCCCCAGA TTGCCCCATTG TTCACCCCTC  
3801 GCCCTGCCCT CCTTGCCTT CCACCCCCAC CATCCAGGTG GAGACCCCTGA  
3851 GAAGGACCCCT GGGAGCTCTG GGAATTGGAG GTGACCAAAG GTGTGCCCTG  
3901 TACACAGGCG AGGACCCCTGC ACCTGGATGG GGGTCCCTGT GGGTCAAATT  
3951 GGGGGGAGGT GCTGTGGGAG TAAAATACTG AATATATGAG TTTTCAGTT  
4001 TTGAAAAAAA AAAAAAAA AAAAAAAA

FIG. 50B



63/78

1 GCAGCGCTGCGTCCCTGCTGCGCACGTGGGAAGCCCTGGCCCCGGCCACCCCCGCGATGCC 60  
 1 CGTCGCGACGCAGGACGACGCGTGCACCCCTCGGGACCGGGGCCGGTGGGGCGCTACGG  
  
 a A A L R P A A H V G S P G P G H P R D A -  
 b Q R C V L L R T W E A L A P A T P A M P -  
 c S A A S C C A R G K P W P R P P P R C R -  
  
 61 GCGCGCTCCCCGCTGCCGAGCGTGCCTCCCTGCTGCGCAGCCACTACCGCGAGGTGCT 120  
 61 CGCGCGAGGGCGACGGCTCGGCACCGAGGGACGACGCGTCGGTATGGCGCTCCACGA  
  
 a A R S P L P S R A L P A A Q P L P R G A -  
 b R A P R C R A V R S L L R S H Y R E V L -  
 c A L P A A E P C A P C C A A T T A R C C -  
  
 121 GCCGCTGGCCACGTTCGTGCAGCGCCTGGGGCCCCAGGGCTGGCGGCTGGTGCAGCGCGG 180  
 121 CGGCGACCGGTGCAAGCACGCCGGACCCGGGTCCGACCGCCGACCACGTCGCC  
  
 a A A G H V R A A P G A P G L A A G A A R -  
 b P L A T F V R R L G P Q G W R L V Q R G -  
 c R W P R S C G A W G P R A G G W C S A G -  
  
 181 GGACCCGGCGGCTTCCGGCGNTGGTGGCCCANTGCNTGGTGTGCGTGCCTGGANGN 240  
 181 CCTGGGCCGCGAAAGGCGCGNACCACCGGGTNACGNACCACACGCACGGGACCTNCN  
  
 a G P G G F P R ? G G P ? ? G V R A L G ? -  
 b D P A A F R A ? V A ? C ? V C V P W ? ? -  
 c T R R L S A R W W P ? A W C A C P G ? ? -  
  
 241 ANGCNGCCCCCGCCGCCCTCCCTGCCAGGTGTCCCTGCCTGAANGANCTGGTGGC 300  
 241 TNCCGNCGGGGGCGGGAGGAAGGCGGTCCACAGGACGGACTTNCTNGACCACCG  
  
 a ? A A P R R P L L P P G V L P E ? ? G G -  
 b ? ? P P A A P S F R Q V S C L ? ? L V A -  
 c G ? P P P P P P S A R C P A \* ? ? W W P -  
  
 301 CCGAGTGCTGCANANGCTGTGCGANCGCGCGCGAANAACGTGCTGGCCTTCGGCTCGC 360  
 301 GGCTCACGACGTNTNCGACACGCTNGCGCCGCGCTTTGACGACCGGAAGCCGAAGCG  
  
 a P S A A ? A V R ? R R E ? R A G L R L R -  
 b R V L ? ? L C ? R G A ? N V L A F G F A -  
 c E C C ? ? C A ? A A R ? T C W P S A S R -  
  
 361 GCTGCTGGACGGGCCCAGGGGGCCCCCGAGGCCTTCACCACCAAGCGTGCAGCTA 420  
 361 CGACGACCTGCCCCGGCGCCCCGGGGCTCCGGAAGTGGTGGTGCACGCGTCGAT  
  
 a A A G R G P R G P P R G L H H Q R A Q L -  
 b L L D G A R G G P P E A F T T S V R S Y -  
 c C W T G P A G A P P R P S P P A C A A T -  
  
 421 CCTGCCAACACGGTGACCGACGCAGCTGCAGGGAGCGGGCGTGGGGCTGCTGCG 480  
 421 GGACGGGTTGTGCCACTGGCTGCGTGACGCCCTCGCCCCGACCCCCGACGACGC  
  
 a P A Q H G D R R T A G E R G V G A A A A -  
 b L P N T V T D A L R G S G A W G L L L R -  
 c C P T R \* P T H C G G A G R G G C C C A A T -

FIG. 51A



64/78

a P R G R R R A G S P A G T L R ? ? C A G -  
 b R V G D D V L V H L L A R C A ? F V L V -  
 c A W A T T C W F T C W H A A R ? L C W W -  
  
 541 GGNNTCCAGCTGCGCCTACCACTGTCGGGCCGCGCTGTACCAAGCTCGGCCTGCNAC  
 CCNAGGGTCAACGCGGATGGTNCACACGCCCGCGACATGGTCGAGGCCGACCGNTG 600  
  
 a G S Q L R L P ? V R A A A V P A R R C ? -  
 b ? P S C A Y ? V C G P P L Y Q L G A A T -  
 c ? P A A P T ? C A G R R C T S S A L ? L -  
  
 601 TCAGGCCCGGCCCGCCACACGCTANTGGACCGAANGCGTCTGGGATCCAACGGCCT  
 AGTCCGGGCCGGGGCGGTGTGCGATNACCTGGGCTNCGCAGACCCTAGGTTGCCCGGA 660  
  
 a S G P A P A T R ? W T R ? R L G S N G P -  
 b Q A R P P P H A ? G P E ? V W D P T G L -  
 c R P G P R H T L ? D P ? A S G I Q R A W -  
  
 661 GGAACCATAGCGTCAGGGAGGCCGGGCTCCCTGGCTGCCAGCCCCGGTGCAGAGGAG 720  
 CCTTGGTATCGCAGTCCTCCGGCCCCAGGGGACCCGACGGTCCGGGCCACGCTCCTC  
  
 a G T I A S G R P G S P W A A S P G C E E -  
 b E P \* R Q G G R G P P G L P A P G A R R -  
 c N H S V R E A G V P L G C Q P R V R G G -  
  
 721 GCGCGGGGGCAGTGCCAGCCGAAGTCTGCCGTTGCCCAAGAGGCCAGGGTGGCGCTGC  
 CGCGCCCCCGTCACGGTCGGCTTCAGACGGCAACGGTTCTCCGGTCCGCACCGCGACG 780  
  
 a A R G Q C Q P K S A V A Q E A Q A W R C -  
 b R G G S A S R S L P L P K R P R R G A A -  
 c A G A V P A E V C R C P R G P G V A L P -  
  
 781 CCCTGAGCCGGAGCGGACGCCCGTTGGCAGGGGCTGGGCCACCCGGCAGGACGCC  
 GGGACTCGGCCTCGCCTGCCGGCAACCGTCCCCAGGACCCGGTGGGCCCGTCTCGGG 840  
  
 a P \* A G A D A R W A G V L G P P G Q D A -  
 b P E P E R T P V G Q G S W A H P G R T P -  
 c L S R S G R P L G R G P G P T R A G R L -  
  
 841 TGGACCGAGTGACCGTGGTTCTGTGTGGTGTACCTGCCAGACCCGCCAGAACGCCAC  
 ACCTGGCTCACTGGCACAAAGACACACCACAGTGGACGGTCTGGCGGCTTCGGTG 900  
  
 a W T E \* P W F L C G V T C Q T R R R S H -  
 b G P S D R G F C V V S P A R P A E E A T -  
 c D R V T V V S V W C H L P D P P K K P P -  
  
 901 CTCTTGGAGGGTGCCTCTGGCACCGGCCACTCCACCCATCCGTGGCCGCCAGCA  
 GAGAAACCTCCCACCGCAGAGAGACCGTGCCTGGTAGGCACCCGGTGGCGGTCTCGGT 960  
  
 a L F G G C A L W H A P L P P I R G P P A -  
 b S L E G A L S G T R H S H P S V G R Q H -  
 c L W R V R S L A R A T P T H P W A A S T -  
  
 961 CCACCGGGCCCCCATCCACATCGCGGCCACCACGTCCCTGGGACACGCCCTGTCCCCCG  
 GGTGCGCCGGGGTAGGTGTAGCGCCGGTGGTGCAGGACCCGTGCGGAACAGGGGC 1020

FIG. 51B

+



MAY 03 2004

65/78

FIG. 51C

1



66/78

a S C G P A C A G W C P Q A S G A P G T T -  
 b R A G L P A P A G A P R P L G L Q A Q R -  
 c V R A C L R R L V P P G L W G S R H N E -  
  
 AACGCCGTTCTCAGGAACACCAAGAAGTTCATCTCCCTGGGGAAAGCATGCCAAGCTCT  
 1501 -----+-----+-----+-----+-----+-----+ 1560  
 TTGCGCGAAGGAGTCCTTGTGGTTCTCAAGTAGAGGGACCCCTTCGTACGGTTCGAGA  
  
 a N A A S S G T P R S S S P W G S M P S S -  
 b T P L P Q E H Q E V H L P G E A C Q A L -  
 c R R F L R N T K K F I S L G K H A K L S -  
  
 CGCTGCAGGAGCTGACGTGGAAGATGAGCGTGCCTGGCTCGCAGGAGCC  
 1561 -----+-----+-----+-----+-----+-----+ 1620  
 GCGACGTCCTCGACTGCACCTCTACTCGCACGCCCTGACGCCAACGACGCGTCCTCGG  
  
 a R C R S \* R G R \* A C G T A L G C A G A -  
 b A A G A D V E D E R A G L R L A A Q E P -  
 c L Q E L T W K M S V R D C A W L R R S P -  
  
 CAGGGGTTGGCTGTGTTCCGGCCGAGAGCACCGTCTGCGTGAGGAGATCCTGGCCAAGT  
 1621 -----+-----+-----+-----+-----+-----+ 1680  
 GTCCCCAACCGACACAAGGCCGGCGTCTCGTGGCAGACGCACTCCTCTAGGACCGGTTCA  
  
 a Q G L A V F R P Q S T V C V R R S W P S -  
 b R G W L C S G R R A P S A \* G D P G Q V -  
 c G V G C V P A A E H R L R E E I L A K F -  
  
 TCCTGCACTGGCTGATGAGTGTGTACGTCGAGCTGCTCAGGTCTTCTTATGTCA  
 1681 -----+-----+-----+-----+-----+-----+ 1740  
 AGGACGTGACCGACTACTCACACATGCAGCAGCTCGACGAGTCCAGAAAAGAAAATACAGT  
  
 a S C T G \* \* V C T S S S C S G L S F M S -  
 b P A L A D E C V R R R A A Q V F L L C H -  
 c L H W L M S V Y V V E L L R S F F Y V T -  
  
 CGGAGACCACGTTCAAAAGAACAGGCTCTTCTACCGGAAGAGTGTCTGGAGCAAGT  
 1741 -----+-----+-----+-----+-----+-----+ 1800  
 GCCTCTGGTGCAAAGTTCTGTCCGAGAAAAGATGGCTCTCACAGACCTCGTTCA  
  
 a R R P R F K R T G S F S T G R V S G A S -  
 b G D H V S K E Q A L F L P E E C L E Q V -  
 c E T T F Q K N R L F F Y R K S V W S K L -  
  
 TGCAAAGCATTGGAATCAGACAGCACTTGAAGAGGGTGCAGCTGCCAGCTGCGGAAG  
 1801 -----+-----+-----+-----+-----+-----+ 1860  
 ACGTTTCGTAACCTTAGTCTGTCGTGAACCTCTCCCACGTCGACGCCCTCGACGCCCTC  
  
 a C K A L E S D S T \* R G C S C G S C R K -  
 b A K H W N Q T A L E E G A A A G A V G S -  
 c Q S I G I R Q H L K R V Q L R E L S E A -  
  
 CAGAGGTCAGGCAGCATGGGAAGGCCAGGCCGCCCTGCTGACGTCAGACTCCGTTCA  
 1861 -----+-----+-----+-----+-----+-----+ 1920  
 GTCTCCAGTCGCGTAGGCCCTCGTCCGGCGGGACGACTGCAGGTCTGAGGCGAAGT  
  
 a Q R S G S I G K P G P P C \* R P D S A S -  
 b R G Q A A S G S Q A R P A D V Q T P L H -  
 c E V R Q H R E A R P A L L T S R L R F I -  
  
 TCCCCAAGCCTGACGGGCTGCGGCCGATTGTGAACATGGACTACGTCGTTGGAGGCCAGAA  
 1921 -----+-----+-----+-----+-----+-----+ 1980  
 AGGGGTTCGGACTGCCGACGCCGGCTAACACTTGTACCTGATGCAGCACCCCTCGGTCTT

FIG. 51D

+



67/78

FIG. 51E



68/78

a A R \* G M P S S S S R A P P \* M R P A V -  
 b P A E G C R R H R A E L L P E \* G Q Q W -  
 c P L R D A V V I E Q S S S L N E A S S G -  
  
 2461 GCCTCTTCGACGTCTCCTACGCTTCATGTGCCACCACGCCGTGCGCATCAGGGGCAAGT  
 CGGAGAAGCTGAGAAGGATGCGAAGTACACGGTGGTGCAGCGTAGTCCCCGTCA 2520  
  
 a A S S T S S Y A S C A T T P C A S G A S -  
 b P L R R L P T L H V P P R R A H Q G Q V -  
 c L F D V F L R F M C H H A V R I R G K S -  
  
 2521 CCTACGTCCAGTGCCAGGGATCCCGCAGGGCTCCATCCTCTCCACGCTGCTCTGCAGCC  
 GGATGCAGGTACGGTCCCCTAGGGCGTCCCGAGGTAGGAGAGGTGCGACGAGACGTCGG 2580  
  
 a P T S S A R G S R R A P S S P R C S A A -  
 b L R P V P G D P A G L H P L H A A L Q P -  
 c Y V Q C Q G I P Q G S I L S T L L C S L -  
  
 2581 TGTGCTACGGCGACATGGAGAACAAAGCTGTTGGGGATTGGCGGGACGGGCTGCTCC  
 ACACGATGCCGCTGTACCTCTTGTGACAAACGCCCTAACGCCGCTGCCGACGAGG 2640  
  
 a C A T A T W R T S C L R G F G G T G C S -  
 b V L R R H G E Q A V C G D S A G R A A P -  
 c C Y G D M E N K L F A G I R R D G L L L -  
  
 2641 TCGCTTGGTGGATGATTCTTGTGACACCTCACCTCACCCACGCCAAAACCTTCC  
 AGCAAACCACTACTAAAGAACAAACCACTGTGGAGTGGAGTGGCTGGCTGGAGG 2700  
  
 a C V W W M I S C W \* H L T S P T R K P S -  
 b A F G G \* F L V G D T S P H P R E N L P -  
 c R L V D D F L L V T P H L T H A K T F L -  
  
 2701 TCAGGACCTGGTCCGAGGTGTCCCTGAGTATGGCTGCGTGGTAACCTGCGGAAGACAG  
 AGTCCCTGGGACCAGGCTCCACAGGGACTCATACCGACGCACCCTGAACGCCCTCTGTC 2760  
  
 a S G P W S E V S L S M A A W \* T C G R Q -  
 b Q D P G P R C P \* V W L R G E L A E D S -  
 c R T L V R G V P E Y G C V V N L R K T V -  
  
 2761 TGGTGAACCTCCCTGTAGAACGAGGCCCTGGTGGCACGGCTTTGTTAGATGCCGG  
 ACCACTTGAAAGGGACATCTCTGCTCCGGACCCACCGTGCAGAAACAGTCTACGCC 2820  
  
 a W \* T S L \* K T R P W V A R L L F R C R -  
 b G E L P C R R R G P G W H G F C S D A G -  
 c V N F P V E D E A L G G T A F V Q M P A -  
  
 2821 CCCACGGCCTATTCCCTGGTGCAGGCCCTGCTGGATACCCGGACCCCTGGAGGTGCAGA  
 GGGTGCAGGATAAGGGACACGCCGGACGACGACCTATGGCCTGGACCTCCACGTCT 2880  
  
 a P T A Y S P G A A C C C W I P G P W R C R -  
 b P R P I P L V R P A A G Y P D P G G A E -  
 c H G L F P W C G L L L D T R T L E V Q S -  
  
 2881 GCGACTACTCCAGCTATGCCGGACCTCCATCAGAGCCAGTCTCACCTCAACCGCGCT  
 CGCTGATGAGGTGAGAAGTGGAGGTAGTCTCGGTCAAGAGTGGAAAGTTGGCGCCGA 2940

FIG. 51F

+



MAY 03 2004

69/78

FIG. 51G



70/78

a V G S S R G R R \* L P W R P Q P T R H C -  
 b S E A P G D D A D C P G G R S Q P G T A -  
 c R K L P G T T L T A L E A A A N P A L P -  
  
 CCTCAGACTTCAGACCATCCTGGACTGATGCCACCGCCCACAGCCAGGCCAGAGCA  
 3421 -----+-----+-----+-----+-----+-----+ 3480  
 GGAGTCTGAAGTTCTGGTAGGACCTGACTACCGGTGGCGGGTGTGGTCCGGCTCTCGT  
  
 a P Q T S R P S W T D G H P P T A R P R A -  
 b L R L Q D H P G L M A T R P Q P G R E Q -  
 c S D F K T I L D \* W P P A H S Q A E S R -  
  
 GACACCAGCAGCCCTGTCACGCCGGCTCTACGTCCCAGGGAGGGAGGGCGGCCACAC  
 3481 -----+-----+-----+-----+-----+-----+ 3540  
 CTGTGGTCTGGACAGTGGCGCCGAGATGCAGGGTCCCTCCCTCCCCGCCGGTGTG  
  
 a D T S S P V T P G S T S Q G G R G G P H -  
 b T P A A L S R R A L R P R E G G A A H T -  
 c H Q Q P C H A G L Y V P G R E G R P T P -  
  
 CCAGGCCCGACCGCTGGGAGTCTGAGGCCCTGAGTGAGTGTGAGTTGGCGAGGCCCTGCATGT  
 3541 -----+-----+-----+-----+-----+-----+ 3600  
 GGTCGGGCGTGGCGACCCCTCAGACTCCGACTCACTCACAAACCGCTCCGACGTACA  
  
 a P G P H R W E S E A \* V S V W P R P A C -  
 b Q A R T A G S L R P E \* V F G R G L H V -  
 c R P A P L G V \* G L S E C L A E A C M S -  
  
 CCGGCTGAAGGCTGAGTGTCCGGCTGAGGCCCTGAGCGAGTGTCCAGCCAAGGGCTGAGTG  
 3601 -----+-----+-----+-----+-----+-----+ 3660  
 GGCGGACTTCCGACTCACAGGCCGACTCCGACTCGTCACAGGTCGGTCCCGACTCAC  
  
 a P A E G \* V S G \* G L S E C P A K G \* V -  
 b R L K A E C P A E A \* A S V Q P R A E C -  
 c G \* R L S V R L R P E R V S S Q G L S V -  
  
 TCCAGCACACCTGCCGTCTTCACTTCCCCACAGGCTGGCGCTGGCTCCACCCAGGCC  
 3661 -----+-----+-----+-----+-----+-----+ 3720  
 AGGTCTGTGGACGGCAGAAGTGAAGGGGTGTCCGACCGCGAGCCGAGGTGGGTCCCG  
  
 a S S T P A V F T S P Q A G A R L H P R A -  
 b P A H L P S S L P H R L A L G S T P G P -  
 c Q H T C R L H F P T G W R S A P P Q G Q -  
  
 AGCTTTCCCTACCAGGAGCCGGCTTCACTCCCCACATAGGAATAGTCATCCCCAGA  
 3721 -----+-----+-----+-----+-----+-----+ 3780  
 TCGAAAAGGAGTGGTCTCGGGCGAAGGTGAGGGGTGTATCCTTATCAGGTAGGGTCT  
  
 a S F S S P G A R L P L P T \* E \* S I P R -  
 b A F P H Q E P G F H S P H R N S P S P D -  
 c L F L T R S P A S T P H I G I V H P Q I -  
  
 TTGCCATTGTTCACCCCTGCCCTGCCCTCCCTGCCTCCACCCACCATCCAGGTG  
 3781 -----+-----+-----+-----+-----+-----+ 3840  
 AAGCGGTAACAAGTGGGAGCGGGACGGGAGGAACGGAAGGTGGGGTGGTAGGTCCAC  
  
 a F A I V H P S P C P P L P S T P T I Q V -  
 b S P L F T P R P A L L C L P P P P S R W -  
 c R H C S P L A L P S F A F H P H H P G G -  
  
 GAGACCCCTGAGAAGGACCCCTGGGAGCTCTGGGAATTGGAGTGAACCAAAGGTGTGCCCTG  
 3841 -----+-----+-----+-----+-----+-----+ 3900  
 CTCTGGGACTCTCCTGGGACCCCTCGAGACCCCTAAACCTCACTGGTTCCACACGGGAC

FIG. 51H

+



71/78

a E T L R R T L G A L G I W S D Q R C A L -  
 b R P \* E G P W E L W E F G V T K G V P C -  
 c D P E K D P G S S G N L E \* P K V C P V -  
  
 TACACAGGCCAGGACCCCTGCACCTGGATGGGGTCCCTGGGTCAAATGGGGGAGGT  
 3901 -----+-----+-----+-----+-----+-----+ 3960  
 ATGTGTCCGCTCCTGGACGTGGACCTACCCCCAGGGACACCCAGTTAACCCCCCTCCA  
  
 a Y T G E D P A P G W G S L W V K L G G G -  
 b T Q A R T L H L D G G P C G S N W G E V -  
 c H R R G P C T W M G V P V G Q I G G R C -  
  
 GCTGTGGGAGTAAAATACTGAATATATGAGTTTCAGTTTGAAAAA  
 3961 -----+-----+-----+-----+-----+ 4020  
 CGACACCCCTCATTTATGACTTATACTCAAAAGTCAAAACTTTTTTTTTT  
  
 a A V G V K Y \* I Y E F F S F E K K K K K -  
 b L W E \* N T E Y M S F S V L K K K K K K -  
 c C G S K I L N I \* V F Q F \* K K K K K K -  
  
 AAAAAAAA  
 4021 ----- 4029  
 TTTTTTTT

FIG. 51

+



72/78

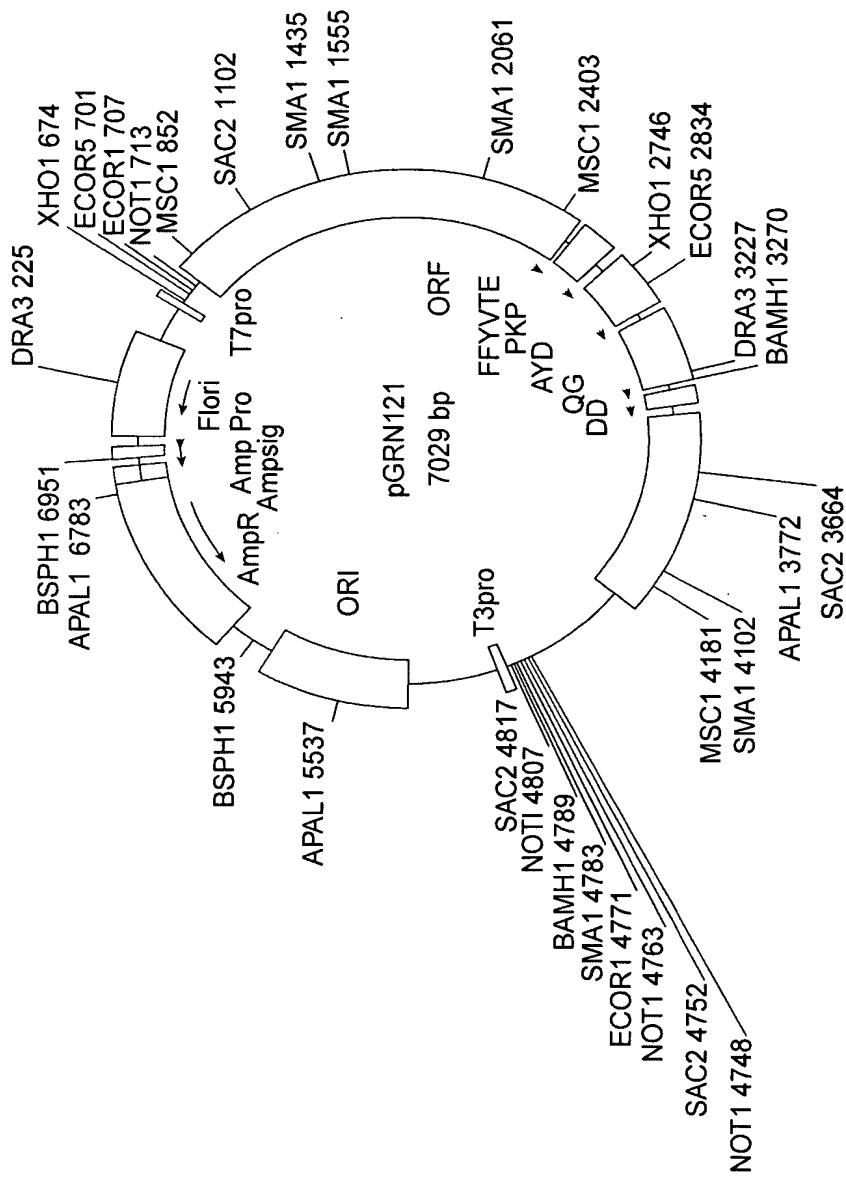


FIG. 52

+



1  
met

GCAGCGCTGCGTCCTGCTGCGCACGTGGAAAGCCCTGGCCCCGGCCACCCCCGGCG ATG

10

pro arg ala pro arg cys arg ala val arg ser leu leu arg ser  
CCG CGC GCT CCC CGC TGC CGA GCC GTG CGC TCC CTG CTG CGC AGC

20

his tyr arg glu val leu pro leu ala thr phe val arg arg leu  
CAC TAC CGC GAG GTG CTG CCG CTG GCC ACG TTC GTG CGG CGC CTG

30

gly pro gln gly trp arg leu val gln arg gly asp pro ala ala  
GGG CCC CAG GGC TGG CGG CTG GTG CAG CGC GGG GAC CCG GCG GCT

40

phe arg ala leu val ala gln cys leu val cys val pro trp asp  
TTC CGC GCG CTG GTG GCC CAG TGC CTG GTG TGC GTG CCC TGG GAC

50

ala arg pro pro pro ala ala pro ser phe arg gln val ser cys  
GCA CGG CCG CCC GCC CCC TCC TTC CGC CAG GTG TCC TGC

60

leu lys glu leu val ala arg val leu gln arg leu cys glu arg  
CTG AAG GAG CTG GTG GCC CGA GTG CTG CAG AGG CTG TGC GAG CGC

70

gly ala lys asn val leu ala phe gly phe ala leu leu asp gly  
GGC GCG AAG AAC GTG CTG GCC TTC GGC TCG CTG GAC GGG

80

ala arg gly gly pro pro glu ala phe thr thr ser val arg ser  
GCC CGC GGG GGC CCC CCC GAG GCC TTC ACC ACC AGC GTG CGC AGC

90

tyr leu pro asn thr val thr asp ala leu arg gly ser gly ala  
TAC CTG CCC AAC ACG GTG ACC GAC GCA CTG CGG GGG AGC GGG GCG

100

trp gly leu leu leu arg arg val gly asp asp val leu val his  
TGG GGG CTG CTG CGC CGC GTG GGC GAC GAC GTG CTG GTT CAC

110

leu leu ala arg cys ala leu phe val leu val ala pro ser cys  
CTG CTG GCA CGC TGC GCG CTC TTT GTG CTG GTG GCT CCC AGC TGC

120

ala tyr gln val cys gly pro pro leu tyr gln leu gly ala ala  
GCC TAC CAG GTG TGC GGG CCG CCG CTG TAC CAG CTC GGC GCT GCC

130

thr gln ala arg pro pro pro his ala ser gly pro arg arg arg  
ACT CAG GCC CGG CCC CCG CCA CAC GCT AGT GGA CCC CGA AGG CGT

140

150

160

170

180

190



74/78

200 210  
 leu gly cys glu arg ala trp asn his ser val arg glu ala gly  
 CTG GGA TGC GAA CGG GCC TGG AAC CAT AGC GTC AGG GAG GCC GGG  
  
 220 230  
 val pro leu gly leu pro ala pro gly ala arg arg arg gly gly  
 GTC CCC CTG GGC CTG CCA GCC CCG GGT GCG AGG AGG CGC GGG GGC  
  
 230 240  
 ser ala ser arg ser leu pro leu pro lys arg pro arg arg gly  
 AGT GCC AGC CGA AGT CTG CCG TTG CCC AAG AGG CCC AGG CGT GGC  
  
 250 260  
 ala ala pro glu pro glu arg thr pro val gly gln gly ser trp  
 GCT GCC CCT GAG CCG GAG CGG ACG CCC GTT GGG CAG GGG TCC TGG  
  
 260 270  
 ala his pro gly arg thr arg gly pro ser asp arg gly phe cys  
 GCC CAC CCG GGC AGG ACG CGT GGA CCG AGT GAC CGT GGT TTC TGT  
  
 280 290  
 val val ser pro ala arg pro ala glu glu ala thr ser leu glu  
 GTG GTG TCA CCT GCC AGA CCC GCC GAA GAA GCC ACC TCT TTG GAG  
  
 290 300  
 gly ala leu ser gly thr arg his ser his pro ser val gly arg  
 GGT GCG CTC TCT GGC ACG CGC CAC TCC CAC CCA TCC GTG GGC CGC  
  
 310 320  
 gln his his ala gly pro pro ser thr ser arg pro pro arg pro  
 CAG CAC GCG GGC CCC CCA TCC ACA TCG CGG CCA CCA CGT CCC  
  
 320 330  
 trp asp thr pro cys pro pro val tyr ala glu thr lys his phe  
 TGG GAC ACG CCT TGT CCC CCG GTG TAC GCC GAG ACC AAG CAC TTC  
  
 340 350  
 leu tyr ser ser gly asp lys glu gln leu arg pro ser phe leu  
 CTC TAC TCC TCA GGC GAC AAG GAG CAG CTG CGG CCC TCC TTC CTA  
  
 350 360  
 leu ser ser leu arg pro ser leu thr gly ala arg arg leu val  
 CTC AGC TCT CTG AGG CCC AGC CTG ACT GGC GCT CGG AGG CTC GTG  
  
 370 380  
 glu thr ile phe leu gly ser arg pro trp met pro gly thr pro  
 GAG ACC ATC TTT CTG GGT TCC AGG CCC TGG ATG CCA GGG ACT CCC  
  
 380 390  
 arg arg leu pro arg leu pro gln arg tyr trp gln met arg pro  
 CGC AGG TTG CCC CGC CTG CCC CAG CGC TAC TGG CAA ATG CGG CCC  
  
 400 410  
 leu phe leu glu leu leu gly asn his ala gln cys pro tyr gly  
 CTG TTT CTG GAG CTG CTT GGG AAC CAC GCG CAG TGC CCC TAC GGG  
  
 410 420  
 val leu leu lys thr his cys pro leu arg ala ala val thr pro  
 GTG CTC CTC AAG ACG CAC TGC CCG CTG CGA GCT GCG GTC ACC CCA

FIG. 53B

+



75/78

430

ala ala gly val cys ala arg glu lys pro gln gly ser val ala  
GCA GCC GGT GTC TGT GCC CGG GAG AAG CCC CAG GGC TCT GTG GCG

440

ala pro glu glu glu asp thr asp pro arg arg leu val gln leu  
GCC CCC GAG GAG GAC ACA GAC CCC CGT CGC CTG GTG CAG CTG

460

leu arg gln his ser ser pro trp gln val tyr gly phe val arg  
CTC CGC CAG CAC AGC CCC TGG CAG GTG TAC GGC TTC GTG CGG

470

ala cys leu arg arg leu val pro pro gly leu trp gly ser arg  
GCC TGC CGC CGG CTG GTG CCC CCA GGC CTC TGG GGC TCC AGG

490

his asn glu arg arg phe leu arg asn thr lys lys phe ile ser  
CAC AAC GAA CGC CGC TTC CTC AGG AAC ACC AAG AAG TTC ATC TCC

500

leu gly lys his ala lys leu ser leu gln glu leu thr trp lys  
CTG GGG AAG CAT GCC AAG CTC TCG CAG GAG CTG ACG TGG AAG

510

met ser val arg asp cys ala trp leu arg arg ser pro gly val  
ATG AGC GTG CGG GAC TGC GCT TGG CTG CGC AGG AGC CCA GGG GTT

530

gly cys val pro ala ala glu his arg leu arg glu glu ile leu  
GGC TGT GTT CCG GCC GCA GAG CAC CGT CTG CGT GAG GAG ATC CTG

550

ala lys phe leu his trp leu met ser val tyr val val glu leu  
GCC AAG TTC CTG CAC TGG CTG ATG AGT GTG TAC GTC GTC GAG CTG

560

leu arg ser phe phe tyr val thr glu thr thr phe gln lys asn  
CTC AGG TCT TTC TTT TAT GTC ACG GAG ACC ACG TTT CAA AAG AAC

580

arg leu phe phe tyr arg lys ser val trp ser lys leu gln ser  
AGG CTC TTT TTC TAC CGG AAG AGT GTC TGG AGC AAG TTG CAA AGC

590

ile gly ile arg gln his leu lys arg val gln leu arg glu leu  
ATT GGA ATC AGA CAG CAC TTG AAG AGG GTG CAG CTG CGG GAG CTG

610

ser glu ala glu val arg gln his arg glu ala arg pro ala leu  
TCG GAA GCA GAG GTC AGG CAG CAT CGG GAA GCC AGG CCC GCC CTG

620

leu thr ser arg leu arg phe ile pro lys pro asp gly leu arg  
CTG ACG TCC AGA CTC CGC TTC ATC CCC AAG CCT GAC GGG CTG CGG

630

pro ile val asn met asp tyr val val gly ala arg thr phe arg  
CCG ATT GTG AAC ATG GAC TAC GTC GTG GGA GCC AGA ACG TTC CGC

640

FIG. 53C





77/78

880  
 leu val thr pro his leu thr his ala lys thr phe leu arg thr  
 TTG GTG ACA CCT CAC CTC ACC CAC GCG AAA ACC TTC CTC AGG ACC  
  
 890 900  
 leu val arg gly val pro glu tyr gly cys val val asn leu arg  
 CTG GTC CGA GGT GTC CCT GAG TAT GGC TGC GTG AAC TTG CGG  
  
 910 920 930  
 lys thr val val asn phe pro val glu asp glu ala leu gly gly  
 AAG ACA GTG GTG AAC TTC CCT GTA GAA GAC GAG GCC CTG GGT GGC  
  
 940  
 thr ala phe val gln met pro ala his gly leu phe pro trp cys  
 ACG GCT TTT GTT CAG ATG CCG GCC CAC GGC CTA TTC CCC TGG TGC  
  
 gly leu leu leu asp thr arg thr leu glu val gln ser asp tyr  
 GGC CTG CTG GAT ACC CGG ACC CTG GAG GTG CAG AGC GAC TAC  
  
 950 960  
 ser ser tyr ala arg thr ser ile arg ala ser leu thr phe asn  
 TCC AGC TAT GCC CGG ACC TCC ATC AGA GCC AGT CTC ACC TTC AAC  
  
 970  
 arg gly phe lys ala gly arg asn met arg arg lys leu phe gly  
 CGC GGC TTC AAG GCT GGG AGG AAC ATG CGT CGC AAA CTC TTT GGG  
  
 980 990  
 val leu arg leu lys cys his ser leu phe leu asp leu gln val  
 GTC TTG CGG CTG AAG TGT CAC AGC CTG TTT CTG GAT TTG CAG GTG  
  
 1000  
 asn ser leu gln thr val cys thr asn ile tyr lys ile leu leu  
 AAC AGC CTC CAG ACG GTG TGC ACC AAC ATC TAC AAG ATC CTC CTG  
  
 1010 1020  
 leu gln ala tyr arg phe his ala cys val leu gln leu pro phe  
 CTG CAG GCG TAC AGG TTT CAC GCA TGT GTG CTG CAG CTC CCA TTT  
  
 1030  
 his gln gln val trp lys asn pro thr phe phe leu arg val ile  
 CAT CAG CAA GTT TGG AAG AAC CCC ACA TTT TTC CTG CGC GTC ATC  
  
 1040 1050  
 ser asp thr ala ser leu cys tyr ser ile leu lys ala lys asn  
 TCT GAC ACG GCC TCC CTC TGC TAC TCC ATC CTG AAA GCC AAG AAC  
  
 1060  
 ala gly met ser leu gly ala lys gly ala ala gly pro leu pro  
 GCA GGG ATG TCG CTG GGG GCC AAG GGC GCC GGC CCT CTG CCC  
  
 1070 1080  
 ser glu ala val gln trp leu cys his gln ala phe leu leu lys  
 TCC GAG GCC GTG CAG TGG CTG TGC CAC CAA GCA TTC CTG CTC AAG  
  
 1090  
 leu thr arg his arg val thr tyr val pro leu leu gly ser leu  
 CTG ACT CGA CAC CGT GTC ACC TAC GTG CCA CTC CTG GGG TCA CTC

FIG. 53E

+



78/78

1100 1110  
 arg thr ala gln thr gln leu ser arg lys leu pro gly thr thr  
 AGG ACA GCC CAG ACG CAG CTG AGT CGG AAG CTC CCG GGG ACG ACG  
 1120  
 leu thr ala leu glu ala ala ala asn pro ala leu pro ser asp  
 CTG ACT GCC CTG GAG GCC GCA GCC AAC CCG GCA CTG CCC TCA GAC  
 1130 1132  
 phe lys thr ile leu asp OP  
 TTC AAG ACC ATC CTG GAC TGA TGGCCACCCGCCACAGCCAGGCCAGAGCAGA  
 CACCAGCAGCCCTGTCACGCCGGCTCTACGTCCCAGGGAGGGAGGGCGGCCACACCC  
 AGGCCCGCACCGCTGGGAGTCTGAGGCCTGAGTGAGTGTGTTGGCCGAGGCCTGCATGTCC  
 GGCTGAAGGCTGAGTGTCCGGCTGAGGCCTGAGCGAGTGTCCAGCCAAGGGCTGAGTGTGTC  
 CAGCACACCTGCCGTCTTCACTCCCCACAGGCTGGCGCTGGCTCCACCCCCAGGGCCAG  
 CTTTTCYTCACCAGGAGCCGGCTTCACTCCCCACATAGGAATAGTCCATCCCCAGATT  
 CGCCATTGTTCACCCYTCGCCCTGCCYTCCTTGCCCTCCACCCCCACCATCCAGGTGGA  
 GACCCTGAGAAGGACCCCTGGGAGCTCTGGAAATTGGAGTGACCAAAGGTGTGCCCTGTA  
 CACAGGCAGGGACCCCTGCACCTGGATGGGGGTCCCTGTGGGTCAAATTGGGGGAGGTGC  
 TGTGGGAGTAAATACTGAATATGAGTTTCAGTTTGRAAAAAAAAAAAAAAAA  
 AAAAAAAA

FIG. 53F

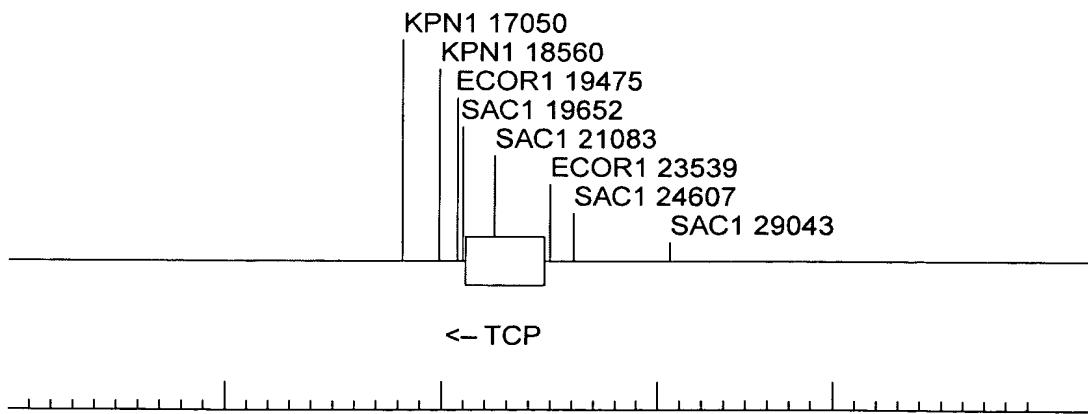


FIG. 54

+